

Journal of Undergraduate Research



JUR Press
Office for Undergraduate Research and Artistry
The Institute for Learning and Teaching
Colorado State University
Campus Delivery 1052
200 W. Lake Street
Fort Collins, CO, 80521-4593
Web: jurpress.org

Designers: Reagan Fitzke
Cover art: Perfect Symmetry – Bryan Bennis

Copyright ©2020 JUR Press. All rights reserved.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the United States Copyright Act without the permission of the copyright owner is unlawful. The copyright of each article is held by the author. Requests for permission or further information should be addressed to the Operations Department of JUR.

ISSN: 2156-5309

Printed in the United States of America

Advertising information: If you are interested in advertising or any other commercial opportunities, please contact the Editor in Chief for JUR at jurpress@gmail.com.

Author Inquiries: For inquiries relating to the submission of articles please visit jurpress.org.

EDITORIAL STAFF

Editor in Chief

Reagan Fitzke

Managing Editor

Zachary Shepard

Associate Editors/Copy Editors

Linda Fischbacher

Cheyenne Izon

Brynn Moore

Raegan Petch

Matthew Yohannes

Tiger Zhang

Affiliate Editors

Saad Bhatti

Sabreen Miah

Monica Zhang

ADVISORY BOARD

Faculty/Staff Advisor

Mary Swanson

Colorado State University

Office for Scholarship and Fellowship Advising

Faculty Advisory Board

Dr. Susan Athey

Business

Dr. Mark Brown

Veterinary Medicine and Biomedical Sciences

Dr. Stephan Weiler

Liberal Arts

Dr. Chris Myrick

Natural Resources

Dr. Matt Hickey

Health and Human Sciences

Dr. Nancy Irlbeck

Agricultural Sciences

Dr. Don Mykles

Natural Sciences

Dr. Tom Siller

Engineering

A Letter from the Editor in Chief

JUR Press is proud to present Volume X of the Journal of Undergraduate Research, our tenth anniversary edition, with published works across a wide range of disciplines by undergraduates from around the world.

We are a journal for undergraduates and by undergraduates. Our slogan – linking the global undergraduate community – captures our mission to serve the interests of worldwide undergraduate thinkers, tinkerers, experimenters, writers, and artists. While JUR does provide a platform for students to publish and showcase their work, we also give students the opportunity to learn about the publication process from start to finish through internships as editors, operations associates, and reviewers. Our undergraduate network has grown to include not only international authors, but also affiliate and satellite editors from around the world! We are truly committed to enhancing the undergraduate experience, and we continually seek to engage as many undergraduate students in our organization as possible.

I am so incredibly thankful to have had the opportunity to serve as Editor-in-Chief for the Journal of Undergraduate Research, and to work alongside a cohort of brilliant editors, authors, and mentors to put together work that honors the experience of novel and innovative ideologies in undergraduate careers. The hard work the staff and authors have poured into this has been truly inspiring. It is essential for this generation to continue to question, learn, and critically think about real-world problems. I am so proud of everyone who has been involved in this process and hope that it is a cause for great celebration to publish works that may have a great impact. I sincerely hope that this work can continue into every individual's graduate and professional careers. Your contributions are significant and will continue to be for much time to come.

I truly hope this journal can continue to empower and embolden the voices of undergraduate scholars long after my tenure is over. It has been inspiring and humbling to commit to this work. I hope to take part of this with me as I continue into my own field of research. Congratulations to all.

Warm regards,

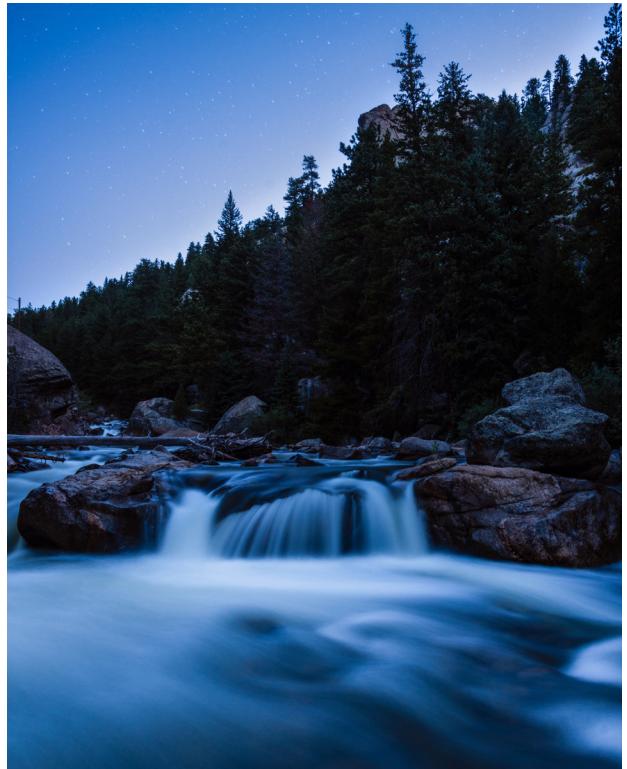
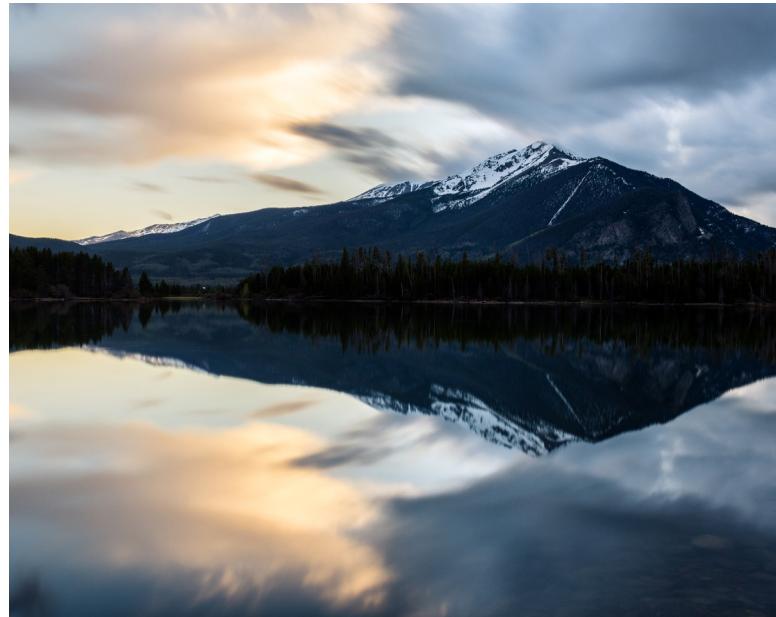
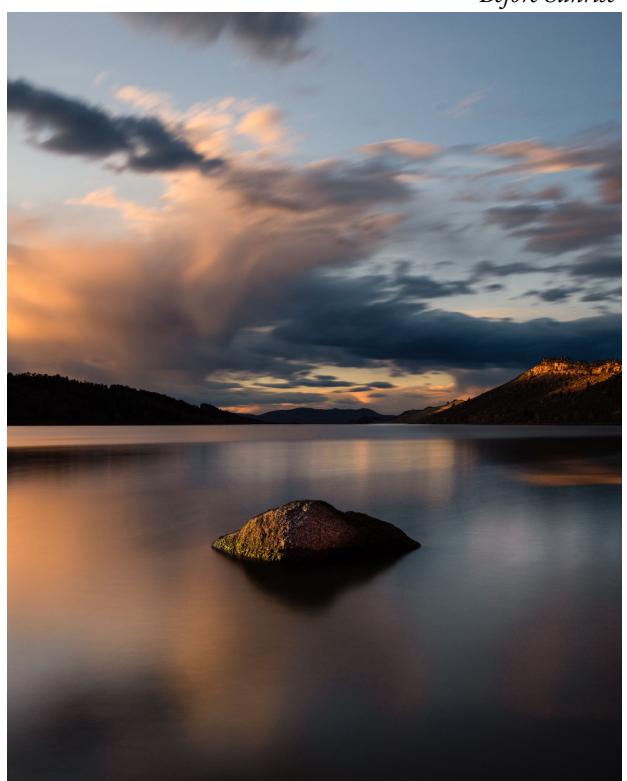
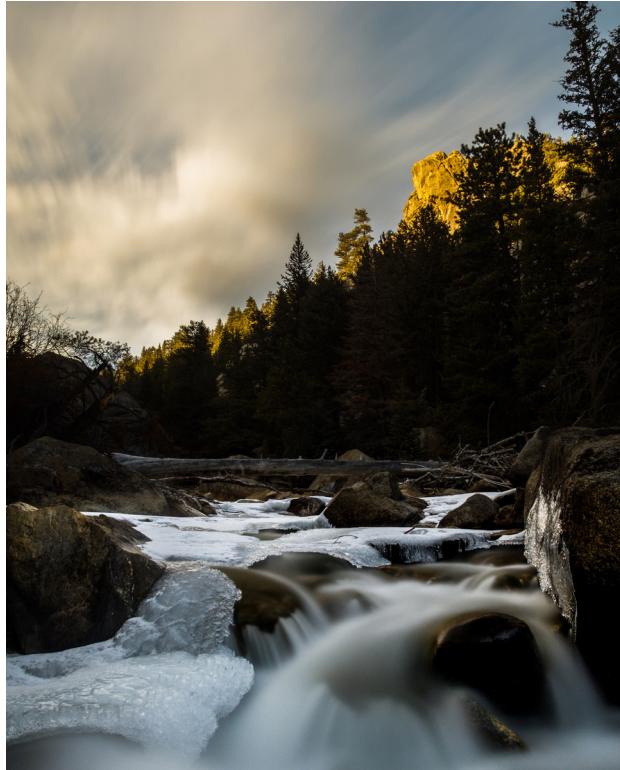
Reagan E Fitzke
Editor in Chief
Journal of Undergraduate Research
JUR-Press

TABLE OF CONTENTS

TITLE	AUTHORS	PAGE
Last Light; Perfect Symmetry; Before Sunrise; Serenity	Bryan Bennis, <i>Colorado State University</i>	1
The Importance of Student Engagement and Experiential Learning in Undergraduate Education (Staff Editorial)	Mark A. Brown, Ph.D, and Reagan E. Fitzke, <i>Colorado State University</i>	2
“Small Fishes in a Big Pond?”: An Exploration on the Current Role of the Community Pharmacist and Local Pharmaceutical Services in England	Laura Minelli, <i>Lancaster University</i>	3
Working Towards a Model of Genetic Profiling for Vulnerability/Resiliency to Sleep Loss through a Summer Research Fellowship in a Military Science Laboratory	Saniya Wadhwa ^{1,3} , Michelle Chen ^{1,3} , Audrey Nikolich ^{2,3} , Ruthie Ratcliffe, MS ³ , and Allison Brager, PhD ³ , ¹ <i>University of Maryland College Park</i> , ² <i>Gettysburg College</i> , ³ <i>Walter Reed Army Institute of Research</i>	14
A Review of Treatments of Substance Use Disorder for Transitional Age Adults	Kai Trepka, Fatima Fairfax, Jeffrey He, Eun Jae Kim, Mingu Kim, Jeongmin Lee, Rajdeep Trilokekar, and Katherine Venturo-Conerly, <i>Harvard University</i>	17
Understanding Benthic Macroinvertebrate Communities in Beaver Meadows of Rocky Mountain National Park	Valerie T. Doebley, <i>Colorado State University</i>	23
Canon and Blues: Diverse Influences on Philip Larkin’s “Aubade”	Brandon Pahl, <i>University of Iowa</i>	33
J.R.R. Tolkien and Environmental Concerns in Mid-20th Century England	Drew Heiderscheidt, <i>University of Northern Colorado</i>	36
“In the Public Interest”: The Historiography of Crusade Finance	Lee Morrison, <i>Florida State University</i>	40

ART

Artwork by Bryan Bennis
Colorado State University



The Importance of Student Engagement and Experiential Learning in Undergraduate Education

BY MARK A. BROWN^{1,2,3,4,5*} AND REAGAN E. FITZKE⁶

COLORADO STATE UNIVERSITY

¹DEPARTMENT OF CLINICAL SCIENCES, COLORADO STATE UNIVERSITY, 80523

²INSTITUTE FOR LEARNING AND TEACHING, COLORADO STATE UNIVERSITY, 80523

³DEPARTMENT OF ETHNIC STUDIES, COLORADO STATE UNIVERSITY, 80523

⁴EPIDEMIOLOGY SECTION, COLORADO SCHOOL OF PUBLIC HEALTH, 80523

⁵CELL AND MOLECULAR BIOLOGY PROGRAM, COLORADO STATE UNIVERSITY, 80523

⁶DEPARTMENT OF PSYCHOLOGY, COLORADO STATE UNIVERSITY, 80523

*CORRESPONDENCE: MARK.BROWN@COLOSTATE.EDU; TEL.: (970) 491-5782

Abstract

Student engagement is among the essential elements in retention programs for undergraduate students¹⁻⁶. Studies illustrate that if students do not perceive a sense of belonging, they are at greater risk of departing their institution prior to completing their degree programs. Recognized as a high impact practice, undergraduate research and other forms of experiential learning have been proven to impart the greatest impacts toward the promotion of student engagement⁷⁻⁸. Herein we underscore the importance of student engagement and experiential learning in undergraduate education.

Introduction

Student engagement¹⁻⁶ and a sense of belonging⁹⁻¹⁴ have been well documented as critical contributors toward successful student retention programs. Higher education research indicates that when students do not perceive a sense of belonging or feel engaged, there is a significantly lower rate of retention. Experiential learning opportunities, such as mentored inquiry, are recognized among the high impact practices with the greatest efficacy among types of programs that seek to engage undergraduates^{7,8}. From improving self-efficacy¹⁵⁻²⁰ to students' sense of belonging⁹⁻¹⁴, experiential learning has been repeatedly proven among positive predictors of academic success, retention, and career success/satisfaction.

Discussion

Experiential learning supports the development of practical problem-solving skills²¹⁻²³. The application of theory, from the classroom, afforded by mentored inquiry improves academic performance and students' ability to define career goals²¹. With experiential learning, undergraduates' regular interface and solutions to real-world issues, followed by contemplation and deliberation related to the results of their decisions, facilitate the growth and development of their understanding associated with their disciplines⁹. Experiential learning is especially appropriate for connecting complex issues across disciplines^{9,10}. The significance of student interactions with faculty mentors as a method of nurturing student engagement through mentored inquiry is being highlighted among best practices for student engagement and retention initiatives¹⁻⁶. Once restricted primarily to STEM and a subset of social science disciplines, experiential learning

and other forms of student engagement are now recognized as essential elements of all undergraduate education. In the past decade experiential learning has been integrated across the breadth of academic disciplines, and the early results are just the tip of the iceberg.

References

- ¹Tinto, V. (2017) "Through the eyes of students." *Journal of College Student Retention: Research, Theory & Practice*. 19.3. Pg 254-269.
- ²Trolian, T., Jach, E., Hanson, J., & Pascarella, E. (2016). "Influencing academic motivation: The effects of student-faculty interaction." *Journal of College Student Development*. 57.7. Pg 810-826.
- ³Komarraju, M., Musulkin, S., & Bhattacharya, G. (2010). "Role of student faculty interactions in developing college students' academic self-concept, motivation, and achievement." *Journal of College Student Development*. 51.3. Pg 332-342.
- ⁴Tinto, V. (2012) *Completing college: Rethinking institutional action*. Chicago: The University of Chicago Press.
- ⁵Upcraft, M., Gardner, J. & Barefoot, D. (Eds.). (2005). *Challenge and support: Creating climates for first year student success*. San Francisco: Jossey-Bass.
- ⁶Pascarella, E., Seifert, T., and Whitt, E. (2008). Effective instruction and college student persistence: Some new evidence. In *The Role of the Classroom in College Student Persistence*. Braxton, J., ed. San Francisco: Jossey-Bass.
- ⁷Kuh, G. (2018). What Really Makes a 'High-Impact' Practice High Impact? Inside Higher Education. May 1, 2018.
- ⁸Kuh, G. (2008, 2011). *Student Success in College: The Promise of High Impact Practices*. Association of American Colleges and Universities.
- ⁹Kuh, G., Kinzie, J., Schuh, J., & Whitt, E. (2005). *Student success in college: Creating conditions that matter*. San Francisco: Jossey-Bass.
- ¹⁰Bergen, J., & Milem, J. (1999). "The role of student involvement and perceptions of integration in a causal model of student persistence." *Research in Higher Education*. 40.6. Pg 641-664.
- ¹¹Stebleton, M., Soria, K., Huesman, R. Jr., & Torres, V. (2014). "Recent immigrant students at research universities: The relationship between campus climate and sense of belonging." *Journal of College Student Development*. 55.2. Pg 196-202.
- ¹²Strayhorn, T. (2008). "Fittin' in: Do diverse interactions with peers affect sense of belonging for black men at predominantly white institutions?" *Journal of Student Affairs Research and Practice*. 45.4. Pg 953-979.
- ¹³Tovar, E. (2013). *A conceptual model on the impact of mattering, sense of belonging, engagement/involvement and socio-academic integrative experiences on community college intent to persist*. Ph.D. Dissertation. Claremont Graduate University, Claremont, CA.
- ¹⁴Sidanis, J., Laar, C., Levin, S., & Sinclair, S. (2004). "Ethnic enclaves and the dynamics of social identity on the college campus: The good, the bad, and the ugly." *Journal of Personality and Social Psychology*. 87.1. Pg 96-110.
- ¹⁵Chemers, M., Hu, L., & Garcia, B. (2001). "Academic self-efficacy and first-year college student performance and adjustment." *Journal of Educational Psychology*. 93.1. Pg 55-64.
- ¹⁶Bong, M. (2001). "Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions." *Contemporary Educational Psychology*. 26.4. Pg 553-570.
- ¹⁷Schunk, D. (1995). "Self-efficacy, motivation, and performance." *Journal of Applied Sport Psychology*. 7.2. Pg 112-137.
- ¹⁸Yueng, M., Brown-Welty, S., & Tracz, S. (2010). "The effects of self-efficacy on academic success of first-generation college sophomore students." *Journal of College Student Development*. 51.1. Pg 50-64.
- ¹⁹Hall, J. M., & Ponton, M. K. (2005). "Mathematics self-efficacy on college freshmen." *Journal of Developmental Education*. 28.3. Pg 26-33.
- ²⁰Gore, P. Jr. (2006). "Academic self-efficacy as a predictor of college outcomes: Two incremental validity studies." *Journal of Career Assessment*. 14.1. Pg 92-111.
- ²¹Boyd, M.K. and Wesemann, J.L. (2009). *Broadening Participation in Undergraduate Research: Fostering Excellence and Enhancing the Impact*. Washington, DC: Council on Undergraduate Research.
- ²²Kardash, C.M. (2000). "Evaluation of an undergraduate research experience: Perceptions of undergraduate interns and their faculty mentors." *Journal of Educational Psychology*. 92.1. Pg 191-201.
- ²³National Research Council. Committee on developments in the Science of Learning and Committee on Learning Research and Educational Practice. (2000). *How People Learn: Brain, Mind, Experience, and School* (expanded ed.) Washington, D.C.: National Academy Press.

“Small Fishes in a Big Pond?”: An Exploration on the Current Role of the Community Pharmacist and Local Pharmaceutical Services in England

BY LAURA MINELLI
LANCASTER UNIVERSITY

1. Introduction

A noticeable discrepancy exists between National Health Service England's (NHSE's) plans for preventive care involving community pharmacy and the public perception of community pharmacy and the local pharmaceutical services (LPSs) it provides. This divergence was noticed by the researcher during her work experience in an English pharmaceutical company implementing LPS in community pharmacies. The research aims to investigate initially the current role of community pharmacists (CPs) in relation to NHSE's proposed integration with general practitioners (GPs). Tuten and Urban's¹ framework in business-to-business (B2B) integration has been utilised to compare the relationship between CPs and GPs and to understand how healthcare professionals' relationships differ from business cooperation. Moreover, the researcher decided to explore the topic of innovation in LPS. The framework developed by Omachonu and Einspruch² has been the basis for understanding healthcare innovation. However, Rogers's idea of diffusion of innovation (1995) has been considered to assess whether LPS follows the same pattern as business-technology-driven innovation. Finally, since LPSs are commissioned to tackle health disparities and to provide patients with wider choice and easily accessible preventive services, this study investigates whether health inequalities are in fact reduced by the presence of LPSs.

To define more precisely this study's focus, the three kinds of clinical services delivered by community pharmacies must be clarified:

- National commissioned services or national services are terms used to refer to those services commissioned by NHSE nationwide.

- Private services are services offered independently by community pharmacies. Thus, these services are not free of charge to the patients since they are not sponsored directly by NHSE.

- Local pharmaceutical commissioned services (LPSs) is the term used to describe the small-scale services commissioned locally by (mainly) local authorities (LAs) and

clinical commissioning groups (CCGs). They are provided in response to local needs, such as the anticoagulation service³. Community pharmacies represent the most accessible point of care, with their presence offering support to the NHS to fight health inequalities.

National Health Service England has recognised the positive effect of allowing community pharmacies to provide the above-mentioned services, and it has developed the pharmacy contractual framework for community pharmacies. The Community pharmacists and their teams provide essential, independent checks and balances within the medicines supply chain, intervening to correct prescribing errors and deal with other issues that could otherwise put patient safety or outcomes at risk. They also use their expertise in medicines procurement to deliver purchasing efficiency, helping the NHS manage the total cost of medicines⁴. The community pharmacy is a major setting for health advocacy in the community⁵, now delivering a wide range of services from the traditional prescription dispensing to clinical services. National Health Services England is assigning major efforts in preventing avoidable illnesses, improving the health of the nation while sustaining the government's financial austerity policy. To do so, more responsibilities are given to the local health government entities (LAs and CCGs), which aim to offer integrated services for communities by understanding their primary clinical needs. In this context, community pharmacies are regarded as the way to reach more patients and to improve their wellbeing.

Research Questions:

The research questions are based around the researcher's personal experience in dealing with the implementation of LPS, and a review of the existing health literature which does not include discussion on the present role of the community pharmacy and LPS. Therefore, these topics have been explored with the guidance of business concepts, which aim to fully uncover the following research questions:

1. What is the current perspective on the role of the Community Pharmacist in

delivering clinical services and the integration with the General Practitioner?

2. In what form is innovation present in Local Pharmaceutical Services (LPSs)?

3. How do Local Pharmaceutical Services (LPSs) support the fight against health inequalities?

In investigating these questions, this study reveals the currently missing integration between healthcare providers, despite NHSE's objectives reported in the “Five-Year Forward View” (5YFV)⁷, in implementing integration and ensuring efficient use of the CP in delivering preventive services. It has been discovered that competition in delivering clinical service is one of the reasons for the lack of a strong relationship between CPs and GPs. The concept of innovation in LPS turned out to be linked to the accessibility of new services by the population rather than the introduction of a completely new service derived by technology. However, “technology” has been used as a key term by the interviewees when addressing the issue of integration among healthcare providers: A new form of communication about patients' health driven by technology is suggested to be employed to ensure effective communication among providers. This solution would also support the development of integration among all healthcare providers. LPSs are commissioned to meet the needs of the population, but their variability in regard to what is offered causes them to have a high number of local commissioning services which are not advertised correctly, leaving the public unaware of what is offered in the pharmacy. This widespread lack of knowledge is a further lost opportunity for expanding preventative measures, since they are underutilised as a result.

2. Context

The United Kingdom health care system is characterised by the only state provider named National Health Service (NHS). The NHS health care system makes citizens recipients of free care, regardless of their socioeconomic status, while simultaneously attributing to them the role of contributors to single-payer insurance⁸. NHSE has

Timeline of Recent NHS England Policies

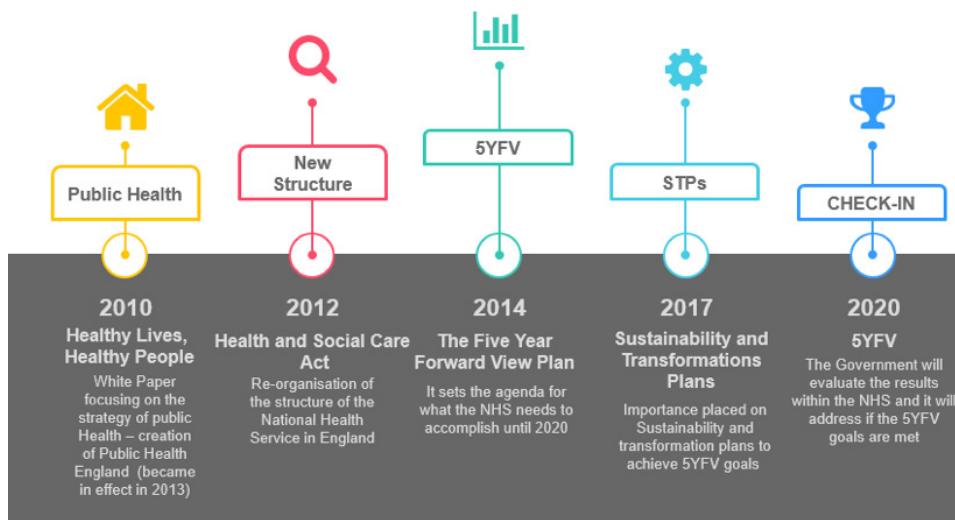


Figure 1. 10-year Timeline of NHS England Policies Changes

encountered many issues, four predominant ones being as follows: financial pressure, an increasingly unhealthy population, rising waiting times, and health inequalities. NHSE was supposed to be a small health care system, but now it has grown, so the short-term financial pressure is a major challenge in NHSE operations. The budget gap increases year over year; it has been forecasted that the gap could deepen to over £10 billion by 2021–2022⁹. The UK is experiencing a huge rise in an ageing and multimorbid¹⁰ population while spending just 8.4% of its GDP to fund NHSE. The Office of National Statistics¹¹ reported 23% potentially avoidable deaths in England and Wales. These deaths, usually linked to heart disease, stroke, and cancers, can be avoided by the patients' pursuing a healthier lifestyle. Just in England, the costs of 'lost productivity from premature mortality and sickness absence resulting from physical inactivity' has been estimated to be £6.5 billion per year¹². Moreover, there were 3.8 million patients waiting for treatment in June 2016, the most patients reported waiting since 2007¹³. Queuing is the control element for care access in NHSE. Without it, everyone would have instant access to care at no cost. Arguably, a lack of restraint would lead the UK population to overuse the health care system, broadening the NHS's financial gap even further. However, the presence of such long waiting times does cause mental and emotional distress for patients¹⁴. Once the element of time is involved, there are accompanying opportunity costs. Improving patients' waiting times leads to better efficiency, thereby reducing negative impacts on patient welfare from queuing and

improving their outcome satisfaction¹⁵. 'One size doesn't fit all'. This simple statement probably best expresses the NHSE's different services, established throughout the counties to deal with health inequalities, which are defined as follows: 'differences between people or groups due to social, geographical, biological factors. These differences have a huge impact because they result in people who are worst off experiencing poorer health and shorter lives'¹⁶.

To overcome the previously mentioned challenges, NHSE has revolutionised its modus operandi, allowing a devolved structure to exist that can effectively meet the clinical needs of a specific area. The NHSE has set up its goals in the 5YFV and plans to receive efficiency savings that will allow it to face the 2020 budget gap by shifting the provision of services to preventive measures and by empowering patients, carers, and communities with their health through accessibility and education about conditions and personalised care budgets. In 2012, the Health and Social Care Act (HSCA)¹⁷ introduced the creation of local bodies across England, called Clinical Commissioning Groups (CCGs). A CCG's role is to improve the health of the area in which they operate while running on efficiency savings in favour of NHSE. The services are commissioned to hospitals, voluntary organisations, and—of central interest to this research—community pharmacies. As NHSE aims to focus on prevention, clinical services in community pharmacies respect NHSE's vision, since they aim to detect illness conditions at an early stage and therefore prevent illness, easing the pressure on GPs and positively reducing the opportunity costs of the GPs.

The Community Pharmacy Forward View¹⁸ proposes three core roles indicating what a CP should be:

1. 'The facilitator of personalised care for people with long-term conditions',
2. "The trusted, convenient first port of call for episodic healthcare advice and treatment',
3. 'The neighbourhood health and wellbeing hub'.

The achievement of the cited goals is an ongoing process supported by the transformation initiatives such as new care and prevention programs developed by NHSE. The Community Pharmacy Forward View is built on the path set by NHSE's 5YFV. The latter argues that change is needed in how healthcare is managed. Preventable illnesses are widespread, and health inequalities are deeply rooted across the country. A new model of care must emerge to ensure that quality care and new and better treatment are available for the population. Thus, pressure on services to meet these demands is building. The primary aim in the 5YFV is prevention, which is an achievable goal if integration of primary care services into the role of CPs is created. NHSE is now working to grant local authorities, CCGs, and local communities more authority and independence in managing the needs of local populations. In this setting, community pharmacies play the critical role of the first point of healthcare advice, where the clinical services delivered aim to prevent illnesses (e.g., flu services), manage chronic conditions (e.g., anticoagulation services), and promote wellbeing (e.g., health check services). In the current context, locally commissioned services ensure that CPs meet the needs of patients by providing well-rounded services. In England, LPSs are free of charge for patients, which should ensure service utilisation and positive health outcomes, such as compliance in disease management in line with NHSE's goals.

3. Review of Relevant Literature

3.1 Developing Integration

Drucker¹⁹ describes difficulties in healthcare management as a 'two-headed monster' to express the idea of conflict between medical and non-medical staff. This idea can be applied to the context of the National Health Service England (NHSE). Since CPs are not directly NHSE's employees, dual complexity in the relationship between GPs and CPs may arise. The two healthcare providers belong to two separate organisations which operate differently, and they offer patients similar clinical services (e.g. Flu vaccination). Therefore, the integration process may result in difficulty if competition is present. For instance, in relation to LPSs,

clinical CCGs and local authorities open the bidding process to any qualified providers (AQP), either pharmacists or doctors, to increase the patients' choices. Hence, the topic of creating business relationships among healthcare providers is explored in relation with the competition that may exist in this environment.

3.2 Healthcare & Business Innovation

Healthcare innovation framework proposed by Omachonu and Einspruch² is presented as lens through which to visualise healthcare innovation. The characteristics of innovation are traced as described in certain business models. In this discussion, the healthcare innovation framework is assessed in contrast to business innovation models. It is argued that healthcare innovation follows similar ideas to those expressed in business innovation models. The same aspects are reconsidered in light of the dissertation's findings, specifically in relation to the clinical services delivered by the CP.

The provision of health care must be driven by innovation to ensure effective care for patients, as planned in NHSE's business objectives. This research considers the presence of innovation in clinical services delivered by the CP and their characteristics against the conceptual ideas of healthcare innovation. Innovation is defined as those changes that help health care practitioners focus on the patient by helping health care professionals work smarter, faster, better, and more cost-effectively²⁰. The healthcare innovation present in clinical services and delivered by CPs can be conceptualised within the framework of healthcare innovation (Figure 2). This framework pictures health care innovation as the result of two external forces, providers and patients, whose needs and capacity of adaptability are the inputs for innovation. However, their role is reciprocal: They are both the catalysts and the recipients of innovation. Innovation must answer questions about how the patients are seen, how they are heard, and how their needs are met.

Taran et al.²¹ stated innovation is the change that occurs in the way businesses operate. Therefore, they analyse the level of change to describe the level of innovation via a three-dimensional approach, first being the innovation's radicality or the level of newness.^c The second dimension comprises the reach of the innovation. So, it poses the question of whether the innovation is new to the company or to the world. The third and final dimension is an innovation's complexity, which is assessed via the change in the organisation's building blocks. The ideal types of innovation are related to their success rates, and success seems to be determined by radicality and reach. On the contrary, complexity does not admit of strong evidence to explain the differences between success and failure. Taran et al.²¹'s research introduced the concept of successful innovation, linked by evidence based on the radicality and reach of the innovation. This evidence aligns with the healthcare innovation framework, as one of the purposes of healthcare innovation is healthcare outreach. In this case, something is defined as healthcare innovation just when it meets the requirement of accessibility. In fact, if innovation is not made commercially feasible for public use, it can't be deemed innovation as it does not present the characteristic of being utilised by the target group who was supposed to enjoy the benefits.

The cycle of innovation described in this chapter follows process improvements rather than disruptive innovation, as often seen in customer goods. In fact, clinical services provided in pharmacies are based upon existing technologies and never undermine the status quo of the firm. Barras²² stated that the cycle of innovation in services is the reverse of the product-cycle theory. Therefore, the first stage of service innovation is based on the objective of increasing the effectiveness of existing service delivery by designing new technologies. In the second stage, technology is applied to improve the quality of service; in the third stage, technology is seen as

the path to reach fully transformed or new services. For example, the anticoagulation management service provided by CPs in England has now introduced novel oral anticoagulants (NOACs), which has meant that routine coagulation monitoring is no longer needed, resulting in lower costs in the long run for anticoagulation management.

4. Methodology

The starting point of this research is the experience gained by the researcher working in pharmacy operations and implementing LPS in the community pharmacy stores of a private company, referred to under the pseudonym PharmaOne. This study is rooted in the current community pharmacy and NHSE context. Thus, by analysing the current insights of experienced professionals who cope daily with contextual changes in the community pharmacy field, the lived experiences of the CPs and their employees can be presented.

Interviews were conducted with 11 members of a private pharmaceutical retail company (PharmaOne) and one Clinical Commissioning Group (CCG's senior leader). The primary goal of this research is to evaluate the role of LPS in the current rapidly changing context of NHSE. The topic is mostly unknown by the public and, due to the complexity of NHSE restructuring, is difficult to interpret without first-hand experienced as guidance. As a result, interviews offered the most appropriate method for exploring this low-profile topic. Therefore, the researcher opted for flexible data categorisation using a Thematic Analysis. The emerging themes identified are listed below and are discussed in Chapter 5:

- the role of the CP and their integration with GPs,
- innovation in LPSs, and
- how healthcare inequalities are tackled by LPS.

5. Analysis and Discussion

The NHSE has recently undergone a devolution process aimed at empowering localities to directly address local needs and increase preventive measures, to ensure a healthier population. Within the local needs context, the topic of innovation present in local pharmaceutical services (LPSs) is explored, along with how health inequalities are addressed. Firstly, the discussion will focus on the integration between CPs and other healthcare providers. An account will be provided of the perspective of CPs of their working realities, as seen through the eyes of the interviewees. Secondly, the concept of innovation as an element present in LPSs will be unpacked about how innovation is present in LPSs and how this affects the population. Finally, the topic of health inequalities will

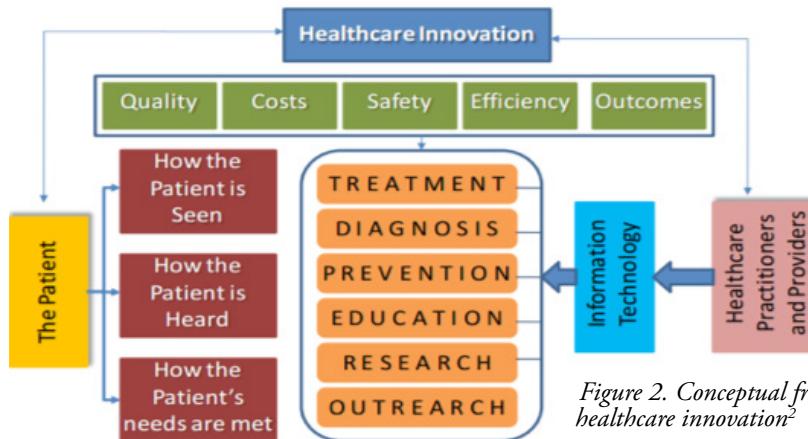


Figure 2. Conceptual framework of healthcare innovation²

be addressed in relation to the role of LPSSs via interviewees' experiences.

5.1 What is the current perspective on the role of the CP in delivering clinical services and the integration with the GP?

As noted in the literature review, the NHSE's plans foresee integration between healthcare providers. There are many advantages to this proposed integration, from reducing service delivery pressure from GPs, to heightened well-being throughout the population via delivery of clinical services in community pharmacies. However, before commenting on provider integration, the role of CPs and how this role supports the NHSE plan in delivering clinical services must be understood. Firstly, the role of the pharmacists will be discussed regarding whether integration is feasible based on their ability to deliver clinical services. A discrepancy has been noted between what CPs can do and the population's perspective of their ability. The presence of integration between healthcare providers, mainly GPs and CPs, will then be assessed via the framework developed by Tuten and Urban¹ to recognise whether healthcare integration follows the same antecedents used to build successful 'business to business' (B2B) cooperation.

5.1.1 People tend to overlook what a community pharmacist can do for them' (Nazli)

Before commencing integration between healthcare providers, such as GPs and CPs, individuals must have a strong confidence in the ability of CPs to deliver clinical services. All interviewees affirmed the ability of the CPs to deliver clinical services. Thus, on the basis of this ability, the NHSE hopes to reduce the pressure that GPs currently face and to simultaneously improve the health of the population by encouraging access to preventive care²³. Increased responsibility is to be given to the CPs via health policy changes, so that they can offer additional clinical services. This action will also lead the CP to be considered a healthcare provider, and CPs will theoretically be included in the healthcare provider integration process. CPs, via the delivery of clinical services, can offer and promote health screening and overall well-being evaluation. They can form the front line to improve the health of the nation, as described by Roberta:

'I think that's where the way community pharmacy is going, and actually a lot of the pharmacies are looking to almost change the way that we're working and offering more clinical services. Pharmacies made that the right thing for the patient. Second, it makes pretty [sic] much more sense for customers. I guess an example is the many community services delivered which are the same as

the ones delivered by GPs, at the moment, and not taking anything away from them'. (Roberta)

In line with Roberta's view, and as expressed in Chapter 2, the role of CPs changed radically, with higher responsibility being placed upon delivering clinical services. Eades, Ferguson and O'Carroll²⁴ concluded that although health policy changes support increased responsibility being given to CPs, the pharmacists themselves were not confident to deliver the clinical services. However, the present study found complete agreement between all interviewees that CPs are confident, willing and ready to deliver clinical services; 'We're working and offering more clinical services' (Roberta). The ideology behind pharmaceutical care symbolises the holistic nature that pharmacies are trying to achieve. The pharmacist's role, rather than being seen as sharply defined, could support services and the overall well-being of the community. This broadening of the CP's role represents what can be considered a sign of 'professional maturity,' stepping into holistic patient. All the interviewees agreed that the pharmacist is able to take on more responsibility in the aftermath of becoming the frontline of prevention and ensure welfare gain:

'We want to support the NHS agenda. Obviously, we are aware of the ability and challenges within the NHS. They're trying to commission some of the services. I think the pharmacies are relying correctly to [offer] support [in] the way they can [...]. I suppose because the pharmacies are open more hours, they are able to reach the population easily. It makes sense for the NHS to commission that services that can be delivered by healthcare professionals. I think at the same time, the pharmacies are really interested in supporting the NHS agenda because community pharmacists have the education and the skills to do so'. (Nazli)

Despite the potential identified in the CP's role, which was shared by all professionals interviewed, the wider population does not hold a similar view. Iversen, Mollison and MacLeod²⁵ reported, regarding the public view of the role of pharmacists, that while customers seem to appreciate the role of the CP towards a healthy-living plan and supporting the work of GPs, they do not seem to fully appreciate the ability of the CP to provide extra services in relation to medicine management. Over 15 years later, and after all the policy changes implemented by the government, the interviewees agreed that a mismatched perception of the CP's role exists:

'I think a lot of it comes down to letting patients know that pharmacists are capable of doing these things. People see that a

pharmacy has different types of medicine, but they don't always realize that they can go there for other things, such as the flu vaccine. I think that it's educating the population: go to your pharmacy for advice and services'. (Leonor)

In summary, a dichotomy exists regarding the CP's role. The evolution of the CP's responsibilities and the changes implemented by the NHSE, both in clinical terms (high numbers of LPSSs commissioned and independent prescribing) as well as in the compensation framework implemented, have allowed the role of CPs to be shifted towards patient care. However, this extension has not been fully appreciated by the customers, as they do not realise what the CP can offer them. It has been postulated that the pharmacy culture must adapt in order to foster the implementation of practice change²⁶. Thus, although the health policies provide CPs with the opportunity to deliver clinical services, customers and clients have not understood this cultural change and believe the CP is mainly a 'medicine dispenser.' Therefore, prior to addressing the topic of integration, a clear issue must be highlighted from the data, that health policy may change but a shift in patients' perception of the true abilities and readiness of the CPs to support their health is needed to fully prepare for integration. The following section will explore the current situation regarding healthcare provider integration and the key conclusions drawn.

5.1.2 'There is a role that is missing to facilitate community pharmacy across primary and secondary care' (Ainhoa)

Healthcare integration is the development and implementation of interconnected working relationships between providers aiming to enhance population health. Schindel et al²⁷ conducted a study based on the perceptions of the pharmacists in the eye of the community after an extended role, in terms of clinical services delivered, had been implemented. The public embraced the benefits in relation to the CP's ability to perform them due to legislative changes. As discussed in the previous section, the legislative changes desired by the NHSE aim towards healthcare-provider integration. Integration is currently a priority in many countries; however, none have developed a comprehensive model. The clinical benefits of integration have already been studied, from palliative service care²⁸ to the successful management of dyslipidaemia²⁹. Examining the concept of integration purely from the business perspective, it was evaluated whether NHSE healthcare integration was created following the same pathway used for B2B cooperation. Tuten and Urban¹ identified the antecedents of a successful B2B

Tuten and Urban (2001)	Interviewees	
Desire for lower costs, e.g., reductions in duplication of effort, paperwork and inefficiency	<p>'One way to avoid that higher cost is to involve pharmacy more but because of that it means that some of those role boundaries start to blur and GPs are independent contractors to the NHS and pharmacies in the main are mostly run as corporate bodies or private individuals.' (Matthew)</p> <p>'A lot of services are provided in hospital or provided by a GP surgery, which actually could be provided by community pharmacy. Transforming the way that you provide services and making sure these services are provided where the patients need them, but also by a professional that can provide it, as opposed to getting the most expensive professional to provide it, will free up money for community pharmacy and the services. It's about shifting where services are provided from' (Ainhoa)</p>	✓
Providing increased service—such as meeting customer needs better and increasing convenience	<p>'Pharmacy has been able to show what it can do in terms of being closer to the patient, quicker to respond to health conditions and like doing a bit more around the prevention of the flu vaccination' (Jude)</p> <p>'I think, well not think, I mean the integration across the whole of the healthcare providers is key, because you can't just in isolation look after an individual. Community pharmacy has probably.... limited access to understanding the whole healthcare needs of that individual. It's actually having that collaborative approach with those healthcare professionals, we need to refer as appropriate and [provide] support for the appropriate' (Stella)</p>	✓
Improving performance indicators—including sales, market shares, and profitability	<p>'When you are delivering service x...there's a commercial reason to do it in a pharmacy you see the till. In a GP's surgery, you don't see the till' (Jack)</p> <p>'Pharmacists would be looking to increase their commission services, they get remunerated to offer them' (Ainhoa)</p>	✓
Increasing product/service quality	<p>'We also have a local enhanced service for supplying quite unusual drugs for people who are dying, palliative care services. When patients are in their last few days or weeks of life, they need some really quite strong painkillers and some sickness medication and these are products that are normally stocked in pharmacies' (Ben -NHSE)</p> <p>'There are some services that a pharmacist can do that doctors are doing a lot more of or anyone else in the practices like flu or like anticoagulant, we can do that a lot easier and quicker in our pharmacies. We've proven that and we've got evidence on that and how it's so much easier, more accessible, and frees up the GP's time and actually, in that case, it does make a lot more sense for some clinical services to be done in pharmacy' (Roberta)</p>	✓
Gaining various benefits of a relationship with a partner—synergy between firms, and a trustworthy partner.	<p>'The GP understands, where the pharmacists can really help them and support them with their patients, the GPs are understanding what services are available through the pharmacy like your vaccination service' (Jack)</p> <p>'That really strong relationship is more complacent. You also have the opposite end of the scale where GPs are there, fully aware of that CPs are there, all those other pharmacists that would have always and historically stayed with them. More of their [CPs'] patients are coming to them, they're probably effective the moment that they have come in to their surgery 'cause CPs and GPs exchange information. There was another side to that where it took a lot of pressure off them [GPs] since that time. Although, yes, they get less flu jabs, but also that they have more appointments available that the next other patient could be sitting in the pharmacy. It depends on the education I think the treatment of patients is getting better when there is this integration present' (Leonor)</p>	✓
Enhancing competitive advantage—such as maintaining a discernible edge relative to competition	Not mentioned and not applicable to this study	X

Table 1. Similarities and Differences between business framework elements regarding B2B relationship¹ and interviewees' responses concerning healthcare providers' integration.

relationship (Chapter 3). Table 1 showcases the elements cited by Tuten and Urban¹ necessary to create a successful partnership in parallel with the interviewees' responses when asked about the relationship between GPs and CPs.

Integration leads to the clinical advantage of patient care, in addition to the benefits of efficient saving. Almost all the antecedents discussed by Tuten and Urban¹ were mentioned in regard to the relationship between CPs and GPs. The business advantages of cooperation between providers are clear to experienced professionals. Interestingly, the interviewees did not mention the final antecedent. The reason for this may lie in the characteristics of the questions asked. This research investigated integration within the healthcare market. In a Beveridge system, as with the English example, competition was introduced by the government based on quality, as quality is higher in competitive markets³⁰. Therefore, clinical services (e.g., anticoagulation) will be awarded to the most qualified provider. Thus, gaining a competitive edge when offering clinical services does not fully reflect the business competition characteristic of self-interest. Healthcare in a Beveridge system is seen largely as a right provided by the government rather than goods to be traded. Thus, from a purely theoretical perspective, market competition should not be present in healthcare provider integration, as healthcare providers are working towards the improvement of the country's health.

It was interesting competition between GPs and CP was mentioned. The presence of competition has been suggested to be one of the key obstacles that frustrate integrative development.

'If we can get rid of that competitive nature and integrate the ways of working to make it in a win-win for both'. (Jack)

'You get this sense of tension between the professions that sometimes gets in the way, as everybody's struggling to survive a lot of the funding shortages within the NHS'. (Ainhoa)

[In relation to Any Qualified Provider Contract – LPS] 'Sometimes it could be interpreted as a competitive threat, then that doesn't exactly foster an environment of good relations for the healthcare professionals working together'. (William)

It is evident that the presence of competition between CPs and GPs does affect their working relationship. If tension and threat are experienced, full integration may be a more distant prospect than first thought. Therefore, while the NHSE plan for integration between healthcare providers has been conceived as an effective solution to reduce costs, especially in preventive care,

the reality of competition may undermine this plan.

Two further elements were discovered underlying this competitive feeling. Firstly, competition was presented by the interviewees as a lack of understanding between the support of the CP for the NHSE plan to open up patient choice of provider. The NHSE allows companies to offer clinical services as GPs to open up that patient's choice through enabling a fair way to leave the choice of provider to the patient.

'At the end of the day, it is patient's choice whether they go to a GP, whether they go to a health store, whether they go to a local pharmacy. I think with any clinical service, it is tried to be done on [the basis of] fairness, so actually it's not direct referrals to a pharmacy or direct referrals to hospital or GP, but it's done in a fair way for both the provider as well as the patient'. (Roberta)

'I think any qualified provider [contract] (AQP) allows patient choice. It's a mechanism for providing greater choice to patients for the services they need. By qualifying as an AQP provider in accessing services, you are then able to advertise your services to patients. I guess the GP is providing the service and providing the script, so the patient may feel more loyalty to remain with surgery, and it's up to the other providers to sell in the benefits of why the patient could choose [them]'. (Ainhoa)

Opening up patient choice regarding provider has been seen as a way to encourage providers to be more responsive to patient preferences about how and where health care is delivered³¹. Moreover, allowing patients to choose their providers encourages these same providers to respond in improving their quality in order to remain in the market and attract and retain patients³². Therefore, allowing patient choice is beneficial to patients, providers and to social welfare gained thanks to the quality improvement of service delivery. Indeed, community pharmacy is transforming, but it still sits in the shadows, and is not taken account of in the national plans.

'Right, community pharmacy is—there's somewhat on the periphery of these changes, GP services are included, but the other three independent contractors' services and primary care are currently not included, although we are just at a point now where we're reaching out to them and starting to have conversations with community pharmacy about how we can bring them in'. (Ben)

"I don't think the NHSE team was planned; they fully realized what pharmacy can do for them. I think very much it's still the primary and secondary care and trying to manage everything through or as much

as possible for GPs and the various other ins and outs, secondary care. However, there are some areas that the pharmacy has been able to show what it can do in terms of being closer to the patient, quicker to respond to health conditions and like doing a bit more around the prevention of the flu [with] vaccination'. (Jack)

The interviewees agreed that considering the national plans of the NHSE, community pharmacy potential is being lost since no holistic plan exists to promote its role. When examining the reasons for not including CPs in the wider plans for a primary role, the unhealthy financial state of the NHSE was blamed. The funding priority is the hospitals. To increase the services delivered in pharmacy, a proportion of the budget must be taken from hospital services by, for instance, a diversion of funding that England is unable to afford. Integration may seem an impossible goal to achieve in the short term. However, thanks to insights from an experienced professional working in relation with Manchester Devolution ('Manc Dev'), attaining integration appears to be closer. Manc Dev allows the area of greater Manchester to be in charge of the social and healthcare budget. Due to this responsibility, the commissioner in charge has called for integration between hospitals, GPs and CPs in providing services such as social prescribing and well-being plans. Manc Dev may appear to be the path to efficiently manage integration, offering a way to use underlying assets such as CPs. However, Manc Dev is the only example of healthcare devolution now present in the UK, and it is only now actionable after years of planning. The road to integration may require years; therefore, a short-term solution must be found, especially when Brexit consequences may negatively affect NSHE research and development, as well as its budget³³. A short-term funding solution may be found to overcome the budget gap; however, this may involve a tax increase, which according to the new government plans will solely support NHS operations³⁴.

In conclusion, the data obtained here indicates that the path towards integration is more complex than simple policy implementation. Although the benefits of partnerships are appreciated, the presence of competition between GPs and CPs is an obstacle for strong working relationships. While the NHS is attempting to broaden patient choice by assigning clinical services to CPs, this action may unintentionally lead to competition to obtain patients, rather than appreciation for the benefits in terms of service quality and reduction in delivery pressure from a GPs perspective. Moreover, it has been noted that community pharmacies

remain in a peripheral planning position. Therefore, integration between GPs and CPs may involve a greater use of resources, either money and time, than originally forecasted by the NHSE.

5.2 In which forms is innovation present in LPSs?

In business services, market orientation impacts the innovation process³⁵. This insight also holds true when considering the healthcare environment, as highlighted by the healthcare innovation framework conceptualised by Omachonu and Einspruch². After some consideration of the role of integration, the topic of innovation in clinical services delivered by the CPs was discussed with the interviewees. Interviewees agreed that innovation is present in LPSs and stated the logic behind their answers. It is generally believed that local pharmaceutical services allow innovative treatments to be widely accessible to the population.

'I think locally commissioned service [Local pharmaceutical services] and private services allow for more innovation because they have developed locally. They meet local needs. What you need to remember with the NHS is that it is formed at a local level, and a lot of the care pathways between localities are so different. Rather than changing everything and having a blank sheet of paper, you build care pathways on what's there currently and what the patients are used to, and you try to improve them. Therefore, local commissions and private services can enable you to do that and provide the local flexibility to meet local needs'. (Ainhoa)

Ainhoa's claim is supportive of the conceptual framework of healthcare innovation², since the role of patients is cited as the reason for why innovation is strongly present in private services and in LPSs. In fact, the framework clearly shows that innovation is driven by how the patient is seen, is heard and how his/her needs are met. Moreover, the idea of clinical services provided by a CP following a service improvement path was confirmed: 'You build care pathways on what's there currently and you try to improve them' (Ainhoa). Therefore, the first stage of the reverse cycle of service innovation (see 22 Chapter 3) is fulfilled. The effectiveness of the service is in fact increased by adding improved elements (applying new technologies) to the existing service rather than via a completely disruptive innovation.

'They [LPSs] absolutely are innovative, [...] like self-testing. That I believe, that I'm aware of, is only operated in Bucks field, but as that grows over yet, learns a lot, it may turn out to be the best thing for the patient; it empowers the patient'. (Leonor)

'I think the biggest thing we've seen in regards to innovation is things like our anti-

coag services. In a GP surgery or hospital you have to have a full phlebotomy, blood test, and then the test results get sent off and it could take a couple of days, sometimes maybe even a couple of hours, but [after] a couple of days it comes back, and by then your INR could be over or under and that, [which] could be life-risking. Whereas in pharmacy now, you can go in, have a blood sample taken and get your INR tested there and then within seconds. It's a (sic) 100% proven, so I think that's innovation. I don't think we do enough of it, I think there are so many other things that we could potentially do, whether it's around diabetes or other tests that we can do in community pharmacy that we just haven't gotten quite there yet, but there are other things we can do, and I think anti-coag is a great example of what community pharmacy can achieve. I think independent prescribers in a walk-in clinic in a community pharmacy is an excellent idea, and we'll be doing that, for example, in the anti-coag, but I think we could do it in more. We do offer minor ailment schemes where we can offer antibiotics which is [sic] usually prescribed by a doctor, but you can do it on specific tests. Other services that we do are the throat test and treatment service, which is very new; it was a trial done in London'. (Roberta)

Independent prescribing (IP) and patient empowerment have been mentioned as the most common elements of innovation in LPSs. Empowerment is considered successful in managing long-term conditions in patients with diabetes mellitus, which also positively affects the psychology of the patients³⁶. Empowerment is an effective philosophy, yet few programs implement it. Moreover, the innovative aspect of allowing CPs to be prescribers further integrates with the practical necessity of efficient GP time management and improved patient care. Although IP and patient empowerment are innovative elements, the technology behind innovative services in pharmaceutical services is already present. Innovative elements allow service quality to be improved, completing the second stage of reverse service innovation²². LPSs and private services allow technological drive to be provided to the population due to the accessibility of pharmacies. However, the two types of service appear to follow different diffusion curves based on the health policy upon which they rely.

'We take technology which is already developed, where we improved, where we innovate is when we put it in a pharmacy, [...] The services which are publicly funded always bring big volumes. In flu vaccination or MUR^d in the UK where they are publicly fund community fund services, we saw the number of flu vacs. done in-pharmacy

multiply by 10%, 20%, 30%. Those are people, so it is multiplied by 10 or 20. It has a huge impact when the service is free or paid by the privileged health system; you really see the number of services delivered increase by a significant amount of time in the population. When you move to policy-based service, you have the typical curves of people who are using your service'. (Nathan)

Therefore, publicly-founded services, as they eliminate any financial barrier in accessing the innovation, see high number of customers utilising the service. It can be argued that, from the moment a service is publicly funded, the patients belong to the majority phase, if compared to Rogers' diffusion theory³⁷, with no innovators or early adopters. On the contrary, private services, as they are not free of charge, experience that the patients behave in the same way as the customers described by the diffusion of innovation curve by Rogers.

When the topic of innovation was raised in the context of LPSs, the technological advancement of communication between healthcare providers was highlighted. Local pharmaceutical services, such as the anticoagulation service or health checks, require the two healthcare providers to communicate the health records of the patients. There is currently no digital form of efficient communication between providers except for the proposed EMIS^e solution for GPs. Zerfass and Huck³⁸ argued that communication is a key factor to innovation management promotion. Being a communication promoter would lead to the simplified sharing of new ideas, technologies, products and services with followers.

'So, you have currently a shift in the population in favour of digital changing the channel. So, the pharmacist has to be part of this shift, or they can really improve their position in the community by being part of this move. If they are not, if they don't do the work, we can imagine that it won't be the GPs, it won't be the pharmacists, but it will be the digital platform that will deal with the health of patients. At the moment, I'm not so sure that it will be beneficial for patients'. (Nathan)

'We're going to use them to do some health checks, take blood pressure, maybe do a diabetes test. Currently, the only way for the pharmacy to add that information to the record is to basically send a piece of paper back to the general practice, and that's not helpful to anybody, really. We do need to progress nationally to a point where community pharmacies can not only access the record but can amend the record or update the record'. (Ben – NHSE)

The pressure to deliver clinical services calls for an effective and technology-driven

method of communication. In this era of technological processes, using 'a piece of paper back to the general practice' to update patient records is not only open to breaches of confidentiality, but can also be considered a waste of time for the healthcare professional. Drucker³⁹ advanced the importance of effective information sharing for healthcare professionals, as they work in a high-level information-based field. Despite CPs delivering clinical services, they are not granted access to patients' health records to adjust their care accordingly, ensure that correct patient health-related information is exchanged, or ensure up-to-date records. The desired technological process would support information-sharing and improved communication between CP and GP, reducing inefficiency and driving innovation.

5.3 How do LPSs support the fight against health inequalities?

Health inequalities have been a topic of health policy discussion for over a century with no optimal solution found yet. The Acheson Report⁴⁰ documented the presence of health inequalities in England, which were derived from social inequalities (income, social status). Bambra et al.⁴¹ suggested that social interventions based on promoting the greater well-being of disadvantaged groups may reduce the health gap. LPSs have been contracted to tackle the specific needs of the population, and thus to reduce health inequalities in given communities.

'They tend to find that the local commission services are commissioned in response to a local need, and therefore, by their very nature they're addressing health inequalities. By addressing that need, if you just think of emergency hormonal contraception, that service's very often formation (sic) is in response to high levels of teenage pregnancy in a particular area. To make access to that contraception easier to people, and help contribute to reduce teenage pregnancies, just the standards of very common pharmacy services like that over two years are impacting local health inequality'. (Jack)

The health needs of an area must be understood by the NHSE to allow proper services to be commissioned. For this reason, the devolution action has been significant, as clinical commissioning groups (CCGs) are now in contact with the need of their localities. This has allowed preventive measures to be built that ensure that necessary services are appropriately commissioned. Further, the reasoning behind the concept of LPSs addressing health inequalities is related to the accessibility argument presented by CPs.

'I think one of the key things is—when we talk about inequalities—is the

accessibility, the accessibility of the services, because of the location for pharmacy within, whether it's deprived locations, whether it's here. The different locations, they are literally more accessible to a broader level of the population. That was one of the real key benefits the pharmacy has over [other] settings, whether that's GP practices which are in certain locations, or like literally in certain locations, where you have one large GP practice which actually fulfils the requirements of a significant geographical area. Actually, it may not be convenient for people to get to that, whether it's because of the cost of the travelling, the ease of the travelling, the health issues that they have in getting there, community pharmacy and a larger location just give that accessibility convenience". (Stella)

The accessibility argument referred to by the interviewees is the geographical accessibility to care, which moves further from the health economics access to care hypothesis. This hypothesis, advanced within health economics, relates to the price of accessing healthcare in other countries⁴². As England belongs to a universalised free healthcare system, the access to care hypothesis does not provide a solution to the current health inequalities present in the system. Therefore, the introduction of the geographical accessibility argument in this study may be an additional hypothesis to explore in the fight against health inequality.

Murray's review⁴³ champions the role of the CPs in enhancing the healthcare of the population, mainly due to their easily accessible presence in deprived areas. The accessibility argument in healthcare suggests that health inequalities are determined by differences in access to care. In countries with universal healthcare, such as England, the accessibility argument may not unequivocally apply⁴⁴. However, the presence of CPs in the consideration of health inequalities and access to care has yet to be studied. Recently, the positive results of better care have been shown via the utilisation of LPSs, such as lowering the rate of teenage pregnancy by introducing an emergency contraception service. Despite the paradigms of the accessibility of pharmacies and the role played by LPSs in focusing on local health needs, health disparities persist, and further reasons for these inequalities have been identified. The first factor relates to the financial pressure the NHSE is facing:

[Local pharmaceutical services] will certainly help, but they're not going to be the magic bullet that kills the issues that we face. So, it will need a team-based approach from all parts of the system really. [...] but the investment in well-being services and prevention services has been

shrinking nationally. The changes in 2013 certainly didn't help because public health budget nationally has reduced. What we're grappling with locally is finding ways to use NHS money that's previously been used for care, illness services, and freeing that money up to pay for well-being prevention services. That's the right thing to do, but it does mean that we have to stop doing some things that we currently do. Good examples of that are in prescribing, we've implemented normal policies in last year where we're actually stopping giving certain medications because we've decided they are not a priority and that it upsets people. We deal with complaints frequently about it, but we have made a decision that we're going to dis-invest in some things in order to reinvest in other things, which are of higher priority. That's essentially what commission needs. Commission is about deciding what are our priorities and then put[ting] in contracts and finance and all the support in place so that the priorities that we've decided are the ones that are provided. It actually means stopping doing other things'. (Ben – NHSE)

The financial pressure the NHSE is currently under undermines the commission of preventive services. Simultaneously, the budget will be prioritised to short-term goals, e.g., chronic condition management rather than investing in long-term solutions such as preventive services commissioned to the pharmacies (e.g. LPSs).

"I think this variability—I think they (LPSs) adds costing to the NHS. I think there should be some standard contracts that you can cut and paste clauses [from]. dependent on your specific needs. I think that could be facilitated as a higher level within the NHS. You haven't got people developing contracts up and down the country for the same sort of thing". (Ainhoa)

The inequality's often associated sort of with schooling, education, and those other areas, or possibly the long you have to wait to see your GP; Those are more extreme, I think, than the actual physical access to the services that typically pharmacies do at a local level. I don't think it's -- because though most of the big four or five local services are done in most of the areas because they're done all in a slightly different way. I think it adds more inequality to it because they're not all judging each patient on the same merits, even though the attempt was to remove inequality". (Matthew)

"That can be quite a confusing picture, but just so, that exposes some of the counter-arguments against having lots of different locally commissioned services. Because the public awareness of what is available then can be confused because you can go from one pharmacy to the next, and different services

[are] available. That's where national services come in. It would make a sense to have national services because then the public can really understand what is available through pharmacy and all the pharmacies, and therefore access that more routinely'. (Jude)

In addition to the lack of funding issue, two major issues arose when considering the role of LPS and health inequalities: LPS variability and the lack of public knowledge on LPS. Firstly, LPS responds to the need of the population. However, each commissioning group provides contracts with different eligibility criteria for the patients of an area, which differ from other areas. Therefore, there is contra-logic in the accessibility aspect. Although a service is accessible, if the eligibility criteria are different, the level of care varies, and thus health inequalities may rise even further as Matthew stated 'they are all done in a slightly different way, it adds more inequalities'. Furthermore, this variability may not only raise health inequalities, but the administrative cost from the NHSE as Ainhoa's interview draw attention to. For instance, the Anticoagulation service can have many variations depending on the drugs the CCG commissioned, the eligibility criteria of patients and whether the service would include independent prescribing by CPs. The number of variations implies that for every anticoagulation service commissioned the costs for the NHSE will differ. On the contrary, a homogenous service throughout the nation would mean less administrative costs to maintain as expenses will be kept low.

Secondly, variability in providing services causes a lack of knowledge about which services are provided. The lack of awareness of the services is determined by two factors: the lack of focus on advertising in the community pharmacies and the lack of national coverage. According to Grönroos and Ravald⁴⁵, potential clients gather information about the services, while the organisation exposes itself in a proactive manner via advertising, personal selling efforts. When dealing with customers' services, the two parties actively engage in information sharing. However, this double action may not appear in the LPS field, as there is a lack of awareness of what is offered in the pharmacies.

"NHS Choices website, as long as that's available from the pharmacy point of view that could tell what the national, local services or other services are available from that pharmacy. Does the patient know what to look at that? Probably not. If you were to Google, depending what you Google, 'what services available from my pharmacy?' I don't know what comes up. Or if you thought I've got a backache and you put in 'I've got

a backache, what do I do? I don't know if pharmacy would score in that. So, it's a hard one to answer, hard, because I know where you could look, if I wasn't in this role would I know to look at NHS Choices and things like that? No, I wouldn't". (Matthew)

The first website that interviewees stated they referred to for information was NHS Choices. However, they agreed there is a lack of information in the absence of a national, comprehensive, information-based website on LPSs. Moreover, Grönroos and Ravarld⁴⁵ suggested that in the climate of competitiveness, companies should not just apply traditional service marketing, but rather an interactive marketing function, in which every component of the service process (organisations, sellers, buyers) interacts not just in the selling encounter but also in the consumption process. By allowing information to be found, customers' needs are best met, and a relationship is created with the customers that continues after the moment of purchase. Therefore, creating a website for clinical services offered by CPs would be beneficial to both the NHSE, as it would allow patients to easily access information, and to private retail pharmacies, as it would drive sales.

While NHSE plans are directed towards CPs assuming increased responsibility, there is a lack of confidence by the public regarding the capacity of the CP. Simultaneously, the integration path is challenged by the competition arising between CP and GPs in providing clinical services and by the competitive nature of service delivery. LPSs are seen as introducing elements of innovation, which are characterised by pre-existing technology made easily accessible to the community. Furthermore, the role of technology innovation in relation to providing improved information sharing between CPs and GPs in delivering clinical services is needed. Finally, while LPSs may be contracted to address health inequalities and ensure prevention services, financial pressure and variability may negatively affect the NSHE plan and budget. LPS variability may appear to have the unintended consequence of leaving patients unaware of the available services. Neither CPs nor the NHS seem to have implemented an effective way of communicating which LPSs are available. The following chapter will present a summary of the findings that emerged from the collected data.

6. Conclusion

This research also exposes issues around the gaps in the process of the integration of healthcare providers. Moreover, it shows how LPS may be the right way to tackle health disparities. However, due to NHSE's

financial pressure, and the lack of exposure, LPS may remain an afterthought rather than being at the forefront of illness prevention.

6.1 Core Findings

The interviews made clear that the research questions offered a conversational starting point with industry experts. In fact, the research questions allowed the interviewees to freely share their thoughts and insights on the topic in question, but they also took the time to explore the topics and the current reality in more depth. The data gathered interestingly challenge assumptions on the role of the CP and LPS and show the gaps in the implementation process of the national plans.

'Relationships between the two [CPs and GPs] are hit and miss.' (Matthew)

From the interviews, it is apparent that there is a discrepancy between what CPs can do and what the public perceives they can do. Moreover, while NHS's plan is to integrate the role of healthcare providers to open the path to complete care, the competitiveness of offering clinical services may limit the national plan. As cited by Matthew, CP and GP is a 'hit and miss', meaning that while the two healthcare providers are supposed to cooperate in widening patient's choices, their cooperation is not fully developed nationwide, thus missing the opportunity to provide better care.

'It's not a brand-new innovation that you do, that there is already some proof of the interest in that innovation from a purely pharmaceutical, general perspective.' (Nathan)

The role of innovation in LPS is completely different from the idea of pure novelty usually experienced in other sectors, as clearly stated by Nathan's comment that 'It's not a brand-new innovation'. LPS allows clinical services to be accessible to the population. Patient-empowerment clinical services such as self-testing anticoagulation or LPSs supporting CPs to take on the role of independent prescribers (IPs) have been cited as an example of innovation introduced in the community pharmacy. These two elements were suggested to be more inclusive when the NHSE commission a LPS. Moreover, since enhanced services (e.g. anticoagulation) necessitate the patients' health records to be updated, communication between GPs and CPs is necessary. It has been found that there is no system in place to ensure that the CP tracks patients' health information and health records. It has been noted that technology-driven system encouraging GP and CP communication about health records could not only represent an innovative element for LPS, but could also enhance the integration among the providers:

'LPS will certainly help, but they're not

going to be the magic bullet that kills the issues that we face.' (Ben)

Local pharmaceutical services are commissioned to answer the local need and to reduce health inequalities in specific areas. However, it has been highlighted how their variability may not only increase health inequalities, but also induce confusion to patients. There is no national website clarifying which LPSs are provided where, and which eligibility criteria are included. The most cited website to support health decisions in relation to access to care is NHS Choice. However, this site does not allow a comprehensive view of LPS. Hence, while LPS are directly commissioned to promote health equity, the lack of awareness by the public of what is offered causes a loss of welfare, as patients in needs will not utilise the services. Moreover, the variability among LPS creates different eligibility criteria for patients to access the service, which may result in a rise in inequality.

6.2 Core Contributions

Although earlier research into the role of the CP has been conducted (25,46), the majority of the literature is dated over three years ago. It does not capture the current climate of the community pharmacy. In fact, 2013 marks a devolution action by the NHS, when more responsibility was given to local teams. The commissioning of services to community pharmacies has increased since 2014. Despite the programs of NHSE to consolidate the role of CP as a healthcare advisor, there are issues related to public perception. Moreover, whereas healthcare-provider integration has been considered a priority in healthcare efficiency, it does not appear to be a simple process, as elements of competition arise in the relationships between GPs and CPs when delivering clinical services. The relationship between the two healthcare providers has been analysed via the B2B framework¹. The antecedents described in the framework have been confirmed to be present in the relationship between CPs and GPs, proving that

1. integration is a feasible goal in community healthcare. However, it is necessary to ensure competition is mitigated; and

2. the competition among GPs and CPs can be analysed via business framework. Although, the topic on the research is relevant to the healthcare market, further research to alleviate the competitive nature of their relationship could be explored, conceiving them as two business operating rather than as healthcare providers.

This research offers a small-scale exploration into the role of community pharmacies and local pharmaceutical

services in the currently changing context. This exploration could be potentially useful for policy when looking at the goals to be achieved in relation to prevention and integration. The path to prevention includes:

- patients actively seeing CPs as the first point of advice and service delivery;
- GPs and CPs effectively communicating and exchanging information; and
- accessible and clear information about preventive services.

Despite the NHSE's efforts, the data suggests these objectives have not been achieved. However, further research must be conducted to appreciate fully the issues and current progress toward solving them. It could be useful in performing nationwide independent research aiming to showcase national spending differences in health care, the services provided, and population health outlooks. This action would lead to further transparency in NHSE operations and a clear picture for health policy decision-makers. Since the Dartmouth Atlas project was firstly published in 1988, it has been used as a tool for American health policy, as it acknowledges the fragmentation of the health care system⁴⁷. A similar and consistent approach may be used by the NHSE to support their decisions and policy goals.

Investigation of the role of LPS in the community environment has highlighted the benefits of the geographical accessibility of the services as determinants to reduce health inequalities. This argument introduces the logistic of service providing rather than the costs associated to accessing care, which is one of the hypotheses often mentioned in health economics when exploring health inequalities⁴². The geographical accessibility of LPS has been cited as one of the reasons these services are determined by innovation, as they allow innovation to spread further into the community. The 'outreach aspect' is an element used to describe the process of innovation in healthcare of the both healthcare innovation framework². However, the process of innovation spread follows a different diffusion curve, depending on the health policy it relies upon. Rogers' diffusion theory³⁷ is in compliance with how private services are accessed; on the contrary, LPSs and national services, as they are free of charge, attract patients independently and do not respond to the same diffusion theory as private services.

Finally, this research has been carried out in 2016/2017, allowing it to consider the effect of the devolution action started in 2014 and to explore the recent changes in health care management (e.g. Manchester Devolution⁴⁸). However, 2016 marks also the confirmation of article 50 for Brexit. Brexit will likely affect the way NSH operates,

the funding and the research development available for the United Kingdom⁴⁹. It would be useful to consider whether Brexit will have repercussions and if so, how these repercussions affect the field of clinical commissioned services within pharmacies once negotiation is over and Article 50 is fully implemented.

6.3 Core Recommendations

To ensure providers understand that the delivery of services relates to the advantage of opening patients' choices, there should be double the effort from CCGs and local authorities in partnership with the CP to ensure collaboration is present among the CPs and GPs. The authority of the NHS and the pro-activeness of CPs may lead GPs to entrust CPs and to ensure proper referral of patients to CPs. In this way, patients will recognise the role of the CP in delivering clinical services, and simultaneously, the GPs will be able to effectively see more urgent cases. Nutbeam⁵⁰ showed that the first step towards higher health policy goals, such as prevention programs and patient empowerment, is the transmission of healthcare information. Creating a communal platform allowing the sharing of the availability of services may be the right way to tackle the gap between what is offered and what people know is offered. Moreover, due to the pattern of how we access information, promotion of health prevention programs should be considered, with use of multichannel platforms, such as social media, via companies' pages (e.g. PharmaOne). Furthermore, since NHSE is currently encouraging the use of the summary care record, the same electronic record should be opened to CPs to ensure precise judgement when delivering clinical services and to update patients' information. Interviewees agreed that this particular data-sharing method could improve the effectiveness of communication among healthcare providers, although they also raised concern for patient's privacy and ensuring that the patient understands the role of CP. Data-sharing tools are controversial, and a security policy must be in place before fully developing such a system. However, it has been proven to be extremely effective in Australia⁵¹.

The recommendations would need a double-effort of development and implementation from both private companies, such as PharmaOne, and NHSE. To improve the health and wellbeing of the population, the NHS tries to activate prevention programmes and to place greater responsibility on the CP. However, due to their limited resources, the budget is still streaming to hospitals and GPs. Prevention assumes a peripheral position. To overcome the budget gap, funding is needed. One

possible short-term solution would be to set an "NHSE tax" determined by each CCG. The NHS was originated after World War Two and did not forecast over 60 million people utilising it.

Throughout the research, the potentials of the CP and LPS to tackle the local needs, promote wellbeing and innovation have been shown and supported via business frameworks, which have been found applicable in the healthcare market. However, these potentials are not completely achieved due to lack of a fully integrated healthcare providers' system and the low-profile role of LPS. There is a lack of awareness about what the CP can provide and the accessibility of LPS. Therefore, although CP and LPS could be key players in supporting NHSE's higher goal of prevention, they are shadowed by the role of GPs. Despite NHSE's plans directed towards prevention, efficiency and integration, CPs and LPS are still small fishes in the big pond of the NHS market.

Acknowledgements

I would like to thank all the healthcare experts who volunteered to take part in my research study as you all made me feel incredibly welcome, and your contributions have been invaluable.

I am extremely grateful to the whole JUR team, Katie Brown, Reagan Fitzke and the external reviewers for using their expertise to support the publishing of this research study – thank you for your amazing help!

Footnotes

^aAnticoagulation is the 'process of hindering the clotting of blood'. The anticoagulation services ensure patients with blood-clotting issues to be monitored and managed in the community pharmacy. Around 1.3 million patients are prescribed anticoagulant therapy in the UK⁴.

^bMultimorbid is commonly defined as the presence of two or more chronic medical conditions in an individual, and it can present several challenges in care, particularly with higher numbers of coexisting conditions and related polypharmacy¹⁰.

^cIn the business model developed by Taran et al.²¹, 'radicality' is a critical variable which determines how much an innovation has departed from what was present before.

^dMUR, the Medicine Usage Review, is a nationally-commissioned service in which the CP completes an adherence-centre review on the medicines prescribed to patients with polypharmacy, to determine whether the patient complies with the medication use and that there are no concerns arising from their use.

^eEMIS (Egton Medical Information Systems) is a web system mainly used by physicians to update electronic patient health records. It also allows patients to book GP appointments online and order repeat prescriptions.

^fSummary care record is an electronic health record containing all the clinical information about the care a patient has received.

References

¹Tuten, T. and Urban, D. (2001) "An Expanded Model of Business-to-Business Partnership Formation

and Success." *Industrial Marketing Management* 30.2. Pg 149-164.

¹⁸Omachou, V. and Einspruch, N. (2010) "Innovation in Healthcare Delivery Systems: A Conceptual Framework." *The Innovation Journal: The Public Sector Innovation Journal* 15.1. Pg 1-20.

¹⁹Merriam-Webster.com. (2017). "Medical Definition of ANTICOAGULATION." <<https://www.merriam-webster.com/medical/anticoagulation>> (Accessed 05/03/2017).

²⁰Hammond, R. (2016) "Bridging anticoagulation: perioperative management of patients on anticoagulants." *Clinical Pharmacist* 8.4. <<https://www.pharmaceutical-journal.com/learning/cpd-article/bridging-anticoagulation-perioperative-management-of-patients-on-anticoagulants/20200956.cpdarticle?firstPass=false>> (Accessed 03/03/2017).

²¹Pharmacy Voice. (2016) "Community Pharmacy Forward View." Royal Pharmaceutical Society English Pharmacy Board.

²²Anderson, C. (2000) "Health promotion in community pharmacy: the UK situation." *Patient Education & Counseling* 39.2. Pg 285-291.

²³NHS England. (2014) "Five-Year Forward View (5YFV)." NHS England. Pg 26-38.

²⁴Hills, J., Ditch, John, & Glennerster, Howard. (1994) "Beveridge and social security: An international retrospective." Oxford University Press.

²⁵Roberts, A. (2012) "The funding pressures facing the NHS from 2010/11 to 2021/22 A decade of austerity?." Nuffield Trust. <<http://www.nuffieldtrust.org.uk/publications/decade-austerity-fundingpressures-facing-nhs>> (Accessed 11/14/2016).

²⁶Wallace, E., Salisbury, C., Guthrie, B., Lewis, C., Fahey, T. and Smith, S. (2015) "Managing patients with multimorbidity in primary care." *BMJ* 350.20. Pg 176.

²⁷Office of National Statistic. (2014) "Avoidable mortality in England and Wales:2014.". Gov.UK. <<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/avoidablemortalityinenglandandwales/2014>>. (Accessed 09/28/2016).

²⁸National Institute of Health and Care Excellence (NICE). (2015) "Tackling the causes of premature mortality (early death)." NICE. <<https://www.nice.org.uk/advice/lgb26/chapter/What-can-local-authorities-achieve-by-taking-action-to-reduce-premature-mortality>> (Accessed 09/12/2016).

²⁹Murray, R. Jabbal, J. Thompson, J. Maguire, D. (2016). "Quarterly Monitoring Report." King's Fund. <<http://qmr.kingsfund.org.uk/2016/20/>> (Accessed 11/05/2016)

³⁰Carr, T. Teucher, U. and Casson, A. (2015) "Waiting for scheduled surgery: A complex patient experience." *Journal of Health Psychology*. Pg 1-12.

³¹Barlow, G.L. (2002) "Auditing hospital queuing." *Managerial Auditing Journal* 17.7. Pg 397-403.

³²Nice.org.uk. (2012) "Health inequalities and population health." NICE. <<https://www.nice.org.uk/advice/lgb4/chapter/introduction>> (Accessed 1/01/2016)

³³Gov.uk. (2012) "Health and Social Care Act 2012: fact sheets." Gov.uk. <<https://www.gov.uk/government/publications/health-and-social-care-act-2012-fact-sheets>> (Accessed 02/11/2017).

³⁴Pharmaceutical Services Negotiating Committee (PSNC). (2016) "Community Pharmacy Forward View." PSNC. Pg 1-20.

³⁵Drucker, P. (1976) "Managing the public service institution." *College and Research Libraries* 37. Pg 10.

³⁶Thakur, R., Hsub, S. and Fontenota, G. (2012) "Innovation in healthcare: Issues and future trends." *Journal of Business Research* 65.4. Pg 562-564.

³⁷Taran, Y., Boer, H. and Lindgren, P. (2015) "A Business Model Innovation Typology." *Decision Sciences* 46.2. Pg 301-331.

³⁸Barras, R. (1986) "Towards a theory of innovation in services." *Research Policy* 15.4. Pg 161-173.

³⁹Kelling, S. E., Rondon-Begazo, A., DiPietro Mager, N. A., Murphy, B. L., & Bright, D. R. (2016) "Provision of Clinical Preventive Services by Community Pharmacists." *Preventing chronic disease* 13.E149. Pg 1-7.

⁴⁰Eades CE, Ferguson JS, O'Carroll RE. (2011) "Public health in community pharmacy: a systematic review of pharmacist and consumer views." *BMC Public Health* 11.1. Pg 1-13.

⁴¹Iversen, L., Mollison, J. and MacLeod, T. (2001) "Attitudes of the general public to the expanding role of community pharmacists: a pilot study." *Family Practice* 18.5. Pg 534-536.

⁴²Rosenthal, Meagen M.Breault, Rene R.Austin, ZubinTsuyuki, Ross T. et al. (2011) "Pharmacists' self-perception of their professional role: Insights into community pharmacy culture." *Journal of the American Pharmacists Association* 51.3. Pg. 363 – 367.

⁴³Schindel, T., Yuksel, N., Breault, R., Daniels, J., Varnhagen, S. and Hughes, C. (2017) "Perceptions of pharmacists' roles in the era of expanding scopes of practice." *Research in Social and Administrative Pharmacy* 13.1. Pg 148-161.

⁴⁴Bainbridge, D., Brazil, K., Krueger, P., Ploeg, J., Taniguchi, A. and Darnay, J. (2014) "Measuring horizontal integration among health care providers in the community: an examination of a collaborative process within a palliative care network." *Journal of Interprofessional Care* 29.3. Pg 245-252.

⁴⁵Blum, B., McKenney, J. and Cziraky, M. (2000) "Pharmaceutical Care Services and Results in Project ImPACT: Hyperlipidemia." *Journal of the American Pharmaceutical Association* 40.2. Pg 157-165.

⁴⁶Gaynor, M. (2006) "What Do We Know About Competition and Quality in Health Care Markets?" National Bureau of Economic Research. Working Paper No.12301. Pg 3-28.

⁴⁷Dixon, A., Robertson, R., Appleby, J., Burge, P. and Devlin, N. (2010) "Patient choice". King's Fund.

⁴⁸Hirschman, A. (1970) *Exit, Voice and Loyalty*. Cambridge MA: Harvard University Press.

⁴⁹Mossialos E, Simpkin V, Keown O, Darzi A (2016) "Will the NHS be affected by leaving or remaining in the EU? Briefing." London School of Economics and Political Science website. <www.lse.ac.uk/newsAndMedia/news/archives/2016/06/Leaving-the-EU-poses-critical-threat-to-NHS.aspx> (Accessed 06/30/2016).

⁵⁰Parker, G. (2017) "Lib Dems propose 1 per cent tax increase to fund NHS." Ft.com. <<https://www.ft.com/content/df1fa6e4-324b-11e7-99bd-13beb0903fa3>> (Accessed 05/07/2017).

⁵¹Agarwal, S., Krishna Erramilli, M. and Dev, C. (2003) "Market orientation and performance in service firms: role of innovation". *Journal of Services Marketing* 17.1. Pg 68-82.

⁵²Anderson, R. M., Funnell, M.M., Butler, P.M., Arnold, M.S., Fitzgerald, J.T. and Feste, C.C. (1995) "Patient Empowerment: Results Of A Randomized Controlled Trial". *Diabetes Care* 18.7. Pg. 943-949.

⁵³Rogers, E (1995). *Diffusion of Innovations*. New York: Free Press. Pg 243-251.

⁵⁴Zerfass, A. and Huck, S. (2007) "Innovation, Communication, and Leadership: New Developments in Strategic Communication." *International Journal of Strategic Communication* 1.2. Pg.107-122.

⁵⁵Drucker, P. (1988) "The Coming of the New Organization." *Harvard Business Review*. <<https://hbr.org/1988/01/the-coming-of-the-new-organization>> (05/07/2017).

⁵⁶Acheson, D. (2000) "Independent inquiry into inequalities in health." London: Stationery Office.

⁵⁷Bambra, C., Gibson, M., Sowden, A., Wright, K., Whitehead, M. and Petticrew, M. (2009) "Tackling the wider social determinants of health and health inequalities: evidence from systematic reviews." *Journal of Epidemiology & Community Health* 64.4. Pg. 284-291.

⁵⁸Bhattacharya, J., Hyde, T. and Tu, P. (2014) *Health economics*. Basingstoke: Palgrave Macmillan.

⁵⁹Murray, R. (2016) "Independent community pharmacy services review commissioned by NHS England." *Pharmaceutical Journal*. <<http://www.pharmaceutical-journal.com/news-and-analysis/news-in-brief/independent-community-pharmacy-services-review-commissioned-by-nhs-england/20201095.article>> (Accessed 09/29/2016).

⁶⁰Marmot, M., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., Marmot, M. and Smith, G. (1991) "Health inequalities among British civil servants: the Whitehall II study." *The Lancet* 337.8754. Pg 1387-1393.

⁶¹Grönroos, C. and Ravald, A. (2011) "Service as business logic: implications for value creation and marketing." *Journal of Service Management* 22.1. Pg 5-22.

⁶²Goldstein, R., Hulme, H. And Willits, J. (1998) "Reviewing Repeat Prescribing - General Practitioners and Community Pharmacists Working Together." *International Journal of Pharmacy Practice* 6.2. Pg 60-66.

⁶³Smith, R. (2011) "Dartmouth Atlas of Health Care." *BMJ* 342.23. Pg d1756-d1756.

⁶⁴Greater manchester-ca.gov.uk. (2017). "Health and social care." GMCA. <https://www.greatermanchester-ca.gov.uk/info/20008/health_and_social_care> (Accessed 02/07/2017).

⁶⁵Edwards, N. (2017) "Brexit means... an uncertain future for the NHS?" The Nuffield Trust. <https://www.nuffieldtrust.org.uk/news-item/brexit-means-an-uncertain-future-for-the-nhs?gclid=Cj0KCQjw85DIBRCyARIsALWb159P_8dZa4xa4coragqHapQnp0GJA_LXZjpQ2YGM63y8P17M9RbvA_YaAoPeEALw_wcB> (Accessed 04/29/2017).

⁶⁶Nutbeam, D. (2000) "Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century." *Health Promotion International* 15.3. Pg 259-267.

⁶⁷Manson, C. (2014) "Could controversial data sharing be good for patient health?" The Guardian. <<https://www.theguardian.com/health-care-network/2014/apr/22/controversial-data-sharing-good-patient-health>> (Accessed 04/22/2017)

Working Towards a Model of Genetic Profiling for Vulnerability/Resiliency to Sleep Loss through a Summer Research Fellowship in a Military Science Laboratory

BY SANIYA WADHWA^{1,3}, MICHELLE CHEN^{1,3}, AUDREY NIKOLICH^{2,3}, RUTHIE RATCLIFFE, MS³, AND ALLISON BRAGER, PhD³

¹UNIVERSITY OF MARYLAND COLLEGE PARK, ²GETTYSBURG COLLEGE,

³WALTER REED ARMY INSTITUTE OF RESEARCH

Abstract

Introduction: The following working model came from undergraduate students who participated in the Army Educational Outreach Program (AEOP). AEOP engages, inspires, and attracts the next generation of Science, Technology, Engineering, and Mathematics (STEM) talent by exposing STEM students to research conducted by the Department of Defense (DoD).

Methods and Materials: The students were paired with the Sleep Research Center (SRC) – the largest research center dedicated to the impact of sleep loss on mental acuity in the DoD at the Walter Reed Army Institute of Research (WRAIR). Though WRAIR is best known for developing vaccines for infectious diseases (and carrying on the legacy of Dr. Walter Reed who discovered the vector responsible for yellow fever), groundbreaking work on the functions and substrates of sleep and how sleep enhances mental acuity was also pioneered at WRAIR. Results/Discussion: During their summer fellowship, the students were inspired by two lines of research that WRAIR has pursued. First, WRAIR has developed caffeine dosing schedules used to keep Soldiers awake during combat.² Second, WRAIR has identified new biological targets intended to predict vulnerability/resiliency to sleep loss and subsequent impact on mental acuity.^{1,2} Therefore, the students made it their goal to combine these two lines of research by developing a working model of genetic profiling that could possibly help identify Soldiers whose health and safety are at-risk during combat under compromised sleep.

Introduction

Identifying the problem: lack of sufficient, restorative sleep during continuous combat operations

Seven to nine hours is the gold-standard for sleep optimization of mental acuity, although there is inter-individual variation in this amount.^{1,2} Less than seven hours of sleep per night is strongly linked to lapses in attention, judgement, and emotional reactivity. If insufficient sleep persists, both morbidity and mortality increase.^{1,2,3,4,5,6} Unfortunately, the United States military prides itself on achieving dominance on the battlefield by means of ‘owning the night’ [see current Army Posture Statement]. Even when not operating, Soldiers have constant anticipation of danger and uncertainty. These factors make it nearly impossible for Soldiers to achieve adequate and restorative sleep. Even for Army-employed sleep researchers and physicians tasked with addressing inadequate and non-restorative sleep in the operational environment, the fact remains that they too also suffer from the inability to achieve adequate and restorative sleep during deployment for reasons listed above.⁷

Identifying effective, pharmacological

countermeasures for the problem: caffeine optimization

Caffeine is the most consumed drug in the United States. According to a study by National Institutes of Health (NIH) researchers in which 37,602 individuals completed comprehensive seven-day diaries, 85% of the U.S. population consumed at least one caffeinated beverage per day.⁸ Caffeine consumption by military personnel is higher than the average population and is also the most common stimulant used by military personnel to stay alert and awake in the operational environment.⁹ Caffeine upregulates arousal-promoting (cholinergic) pathways.¹⁰ Caffeine can also modestly delay (also referred to as phase-shift) human sleep/wake and endocrine rhythms.¹¹

The current working model was built specifically on the caffeine research completed at WRAIR. First, WRAIR researchers have developed caffeine dosing schedules using a patented, quick-release caffeinated gum: Military Energy Gum (MEG).¹² One piece of MEG contains 100 mg of caffeine. Kamimori et al. discovered in a dose-dependent clinical trial that the gum is pharmacologically active in less than 10

minutes and that the mechanisms of action were such that caffeine released from the gum quickly crossed the blood-brain barrier via the salivary buccal cavities, bypassing the digestive tract.¹² Second, WRAIR researchers have found that the ability of caffeine to preserve and/or slow the rate of degraded mental acuity with less than seven hours of nighttime sleep is obsolete after three days.¹³ After three days, caffeine cannot substitute the restorative effects of sleep for next-day performance.¹³ The third and most critical finding for the development of the working model is WRAIR’s discovery of large inter-individual variation in responsiveness (e.g., tolerance and sensitivity) to caffeine’s alertness and performance-enhancing effects with sleep loss.^{1,2} This data is critical because it shows that no two people respond similarly to sleep loss, recovery from sleep loss, or the ability of caffeine to mitigate the negative consequences of sleep loss. In summary, these findings make up almost two decades of research on caffeine’s ability to stabilize performance under sleep loss. These findings are also the basis for caffeine dosing schedules adopted by military personnel during deployment as published in *Army Training*

Protocol (ATP) 6-22.5: A Leader's Guide to Soldier Health and Fitness (Figure 1).

Identifying effective, biological countermeasures for the problem: genetic polymorphisms

The third step towards developing a working model to optimize Soldier health and safety was to dissect the genetic landscape of vulnerability and resiliency to sleep loss. Select genes are involved in regulation of sleep amount, sleep timing, and caffeine metabolism.^{14,15,16,17,18} Furthermore, the biochemical actions of caffeine and regulation of sleepiness by adenosine are closely related. Caffeine blocks the release of adenosine, a neurotransmitter that suppresses neural activity in the brain, leading to a desire and biological need to sleep. Adenosine is also a byproduct of wakefulness due to increased production of adenosine triphosphate (ATP).^{16,17} During sleep deprivation, adenosine levels in the brain will continue to rise well beyond normal physiological ranges and will not fall until an individual sleeps.^{16,17}

With this knowledge, the working model was built on six single-nucleotide polymorphisms (SNPs) underlying sleep regulation by adenosine -- ADORA2A, ADA -- and by circadian clock-controlled genes -- PER2, PER3 -- as well as caffeine metabolism at the level of the liver: CYP1A2, and NAT2.^{14,15,16,17,18} Each SNP has a selective predictive role/function pertinent

to sleep regulation, caffeine metabolism, and next-day performance:

1. Adenosine-derived polymorphisms (ADORA2A and ADA) predict next-day performance after normal sleep and/or after sleep loss.^{14,15,19} The particular genetic variants of ADORA2A for this working model were rs5751862, rs5760405, rs2298383, rs3761422, rs2236624, rs35329474, and rs4822492 found on chromosome 22. The particular genetic variants of ADA for this working model were rs73598374 and rs394105 found on chromosome 20. In general, these studies have found that HT4 haplotypes were more resilient to sleep loss than non-HT4 haplotypes. Interestingly, these phenotypes were linked to "genetic trade-offs" such that non-HT4 haplotypes were more sensitive to caffeine compared to HT4 haplotypes, meaning caffeine had the ability to stabilize performance under sleep loss in non-HT4 haplotypes.

2. PER polymorphisms predict preference for early rise/bedtimes (< 0500/ < 2100, EST) or late rise/bed times (> 1000/ > 0100, EST). Preferred rise time/bedtimes are important considerations. First, research has shown that athletes engaged in high-risk physical activity (like Soldiers) perform better in the evening.²⁰ Further, caffeine supplementation can elevate high-risk physical activity performed in the early

morning to levels of (placebo-supplementing) evening performance.²⁰ Second, it has been shown that early risers are more physiologically and psychologically resilient to sleep deprivation.²¹ The particular genetic variants of PER2 for this working model were rs2304672 and rs10462023 found on chromosome 2. The particular genetic variants of PER3 for this working model were rs35426314, rs228669, rs35733104, rs228696, rs35899625, rs228697, and rs17031614 found on chromosome 1. For these particular coding regions, previous studies have found, for example, that PER3 (4/4) genotypes were more resilient to sleep loss compared to PER3 (4/5) genotypes.¹⁹ Unlike research completed with adenosine polymorphisms, the ability of caffeine to stabilize performance in PER3 (4/5) genotyped individuals under sleep loss is still unknown.

3. The two primary polymorphisms conferring inter-individual differences in caffeine metabolism, CYP1A2 and NAT2.^{22,23} The particular genetic variants of CYP1A2 for this working model were rs2069514, rs12720461, and rs762551 found on chromosome 15. The particular genetic variants of NAT2 for this working model were rs1041983 and rs1801280 found on chromosome 8. The selected polymorphisms of CYP1A2 and NAT2 are linked to reduced caffeine sensitivity due to heightened metabolism of caffeine by way of heightened CYP1A2 and NAT2 ratios.²²

Methods and Results

Developing a working model to optimize Soldier health and safety in future studies through understanding of Soldier sleep, caffeine supplementation, and genetic variation in vulnerability/resiliency to sleep loss

The working model is envisioned to be used for future WRAIR- and DoD-directed studies to assist commanders with mission execution and used during the selection process of the Special Operations Command (e.g., Army Rangers). In order for the working model to have direct military application, this would require access to an individual Soldier's blood profile for each genetic variant of ADORA2A, ADA, PER2, PER3, CYP1A2, and NAT2. The four quadrants of the working model are founded on inter-individual variation in vulnerability/resiliency to sleep loss, caffeine sensitivity, and genetic environment (Figure 2). Cognitive performance in the model refers to the ability to maintain stable performance on the psychomotor vigilance test, the gold-standard for assessing real-time mental acuity (reaction time) and alertness under sleep loss.^{1,2,12,13,14}

In the model, A1 individuals (resilient

Table 2-3. Using caffeine under various conditions of sleep deprivation

Condition under which caffeine is used	Guidelines for use
Sustained operations (no sleep)	<ul style="list-style-type: none"> 200 milligrams starting at approximately midnight. 200 milligrams again at 0400 hours and 0800 hours, if needed. Use during daytime hours only if needed.
Night shifts with daytime sleep	<ul style="list-style-type: none"> 200 milligrams starting at beginning of nighttime shift.
Restricted sleep	<ul style="list-style-type: none"> 200 milligrams upon awakening. 200 milligrams again 4 hours later. Discontinue or reduce caffeine intake for the last 4 to 6 hours before initiating sleep.

Figure 1: Adapted from Army Training Protocol (ATP) 6-22.5: A Leader's Guide to Soldier Health and Fitness. Table outlines caffeine dosing schedules adopted by military personnel during deployment. These best practices were developed from over two decades of caffeine research done at the Walter Reed Army Institute of Research (WRAIR) in Washington, D.C.

Cognitive Performance post Sleep Disruption

Caffeine Sensitivity		
	(A) Preserved	(B) Impaired
(1) Sensitive	A1*	B1
(2) Tolerant	A2	B2*

Figure 2: Intra-individual variation in the ability to perform is dependent on two principles. The first principle is intra-individual sensitivity to sleep disruption. Some individuals, due to genetic variants, are more resilient to sleep deprivation, meaning their mental acuity (tested using the psychomotor vigilance test) degrades at a slower rate across sleep loss compared to individuals not (genetically) resilient.^{19,23} The second principle is intra-individual sensitivity to caffeine. Some individuals are highly sensitive to caffeine, allowing for performance enhancement and/or optimization under sleep loss.^{10,13,19}

*to sleep loss and sensitive to caffeine) would be the ideal candidate to perform in military operations. According to our working model, these individuals may have a genetic predisposition that allows them to have stable performance under sleep loss coupled with possibly increased performance- and alertness-enhancing benefits from caffeine. These individuals would be characterized as what military leaders would label as elite performers and 'super Soldiers.' On the other hand, *B2 individuals (vulnerable to sleep loss and tolerant to caffeine) would be least suited to perform military operations* due to their possible decline in performance under sleep loss and an inability of caffeine to subsequently rescue performance to pre-sleep loss levels. Both traits of poor responses to sleep disruption and caffeine could impact military performance during sustained (> 24 h) missions. *A2 (resilient to sleep loss but tolerant to caffeine) and B1 (vulnerable to sleep loss but sensitive to caffeine)* individuals would be moderately suited for military operations. A2 individuals are able to perform normally under sleep disruption, but caffeine would not be able to give these Soldiers a competitive edge like those who fall under A1. B1 individuals would hypothetically be in danger with sleep disruption during military operations, but their high sensitivity to caffeine could hypothetically protect against decreases in military performance under sustained missions. Moreover, this chart helps to classify Soldiers into different categories of hypothetical performance, contributing to the foundations for the development of methods to help Soldiers based on their genetic background.*

Discussion

There were several lessons learned from the students' research fellowships at WRAIR. First, the students learned about the importance of restorative vs. non-restorative sleep during sustained (> 24 hour) military operations. With restorative sleep, individuals can maintain performance. With non-restorative sleep, performance suffers, compromising the health and safety of the Soldiers and the unit. Second, the students learned about the ability of one's genetic background to predict performance under sleep loss that can be rescued, in part, with caffeine supplementation; some individuals cognitively suffer during sleep loss (the B1s and B2s), while other individuals (the A1s and A2s) could be cognitively preserved during sleep loss as measured from gold-standard, real-time measurements of mental acuity (reaction time).

In brief, the working model may help leaders take actions towards removing subjective bias for military selection by

focusing on genetic attributes of elite performers and/or non-performers. The working model may also help select individuals for military service who otherwise would not be considered. In fact, these are the exact intents of the Defense Advanced Research Program Agency (DARPA) program entitled Measuring Biological Amplitude (MBA). There are some limitations with the working model, however.

First, it would require a great deal of knowledge and coordination to align work schedule/mission requirements with knowledge of peak optimal performance predicted based on the genetic variants. Second, while optimization for when a Soldier has to work at a sub-optimal time of the day (e.g., night shift) can be achieved through caffeine supplementation, preventing tolerance to caffeine would still be a concern even with a genetic predisposition for caffeine sensitivity. Finally, even when controlling for an individual's genetic landscape, the fact remains that deployed settings do not permit for restorative sleep and so countermeasures such as frequent napping, early bedtimes, and absence of light/bedside technology (reduces the ability to achieve restorative sleep) must still be considered. To conclude, with greater knowledge of genetic profiling and benefits of restorative sleep for performance, effective countermeasures and strategies could be developed to maximize military performance.

References

¹Belenky, G., Wesensten, N.J., Thorne, D.R., Thomas, M.L., Sing, H.C., Redmond, D.P., Russo, M.B., and Balkin, T.J. (2003) "Patterns of performance degradation and restoration during sleep restriction and subsequent recovery: A sleep dose-response study." *Journal of Sleep Research.* 12. Pg 1-12.

²Rajdev, P., Thorsley, D., Rajaraman, S., Rupp, T.L., Wesensten, N.J., Balkin, T.J., Reifman, J. (2013). "A unified mathematical model to quantify performance impairment for both chronic sleep restriction and total sleep deprivation." *Journal of Theoretical Biology.* 331. Pg 66-77.

³Czeisler, C.A., Duffy, J.F., Shanahan, T.L., Brown, E.N., Mitchell, J.F., Rimmer, D.W., Ronda, J.M., Silva, E.J., Allan, J.S., Emens, J.S., Dijk, D.J., Kronauer, R.E. (1999). "Stability, precision, and near-24- hour period of the human circadian pacemaker." *Science.* 284. Pg 2177-2171.

⁴McCauley, P., Kalachev, L.V., Mollicone, D.J., Banks, S., Dinges, D.F., Van Dongen, H.P. (2013). "Dynamic circadian modulation in a biomathematical model for the effects of sleep and sleep loss on waking neurobehavioral performance." *Sleep.* 36. Pg 1987-1997.

⁵Killgore, W.D., Kahn-Greene, E.T., Lipizzi, E.L., Newman, R.A., Kamimori, G.H., and Balkin, T.J. (2008). "Sleep deprivation reduces perceived emotional intelligence and constructive thinking skills." *Sleep Medicine.* 31. Pg 517-526.

⁶Pace-Schott, E.F., Nave, G., Morgan, A., and Spencer, R. (2012) "Sleep-dependent modulation of affectively guided decision-making." *Journal of Sleep Research.* 21. Pg 30-39.

⁷Good, C., Brager, A., Capaldi, V., and Mysliewic, V. (2020) "Sleep in the United States Military." *Neuropsychopharmacology.* 45. Pg 171-191.

⁸Mitchell, D., Knight, C., Hockenberry, J., Teplansky, R., Hartman, T. (2014). "Beverage caffeine intake in the U.S." *Food and Chemical Toxicology.* 63. Pg 136-142.

⁹Attipoor, S., Manganello, C., Scott, JM., Deuster, PA. (2017). "Usefulness of a risk assessment tool to risk stratify dietary supplements." *Military Medicine.* 182. Pg 2086-2091.

¹⁰Urry, E., Landolt, HP. (2015). "Adenosine, caffeine, and performance: from cognitive neuroscience of sleep to sleep pharmacogenetics." *Current Topics in Behavioral Neuroscience.* 25. Pg 331-366.

¹¹Burke, TM., Markwald, PR., McHill, AW., Chinoy, ED., Snider, JA., Bessman, SC., Jung, CM., O'Neil, JS., Wright, KP Jr. (2015). "Effects of caffeine on the human circadian clock in vivo and in vitro." *Science Translational Medicine.* 305.

¹²Kamimori, GH., Karyekar, CS., Otterstetter, R., Cox, DS., Balkin, TJ., Belenky, GL., Eddington, ND. (2002). "The rate of absorption and relative bioavailability of caffeine administered in chewing gum versus capsules to normal health volunteers." *International Journal of Pharmaceutics.* 234. Pg 159-167.

¹³Doty, T.J., So, C.J., Bergman EM., Trach, SK., Ratcliffe, RH., Yarnell, AM., Capaldi, VF 2nd., Moon, JE., Balkin, TJ., Quartana, PJ. (2017). "Limited efficacy of caffeine and recovery costs during and following 5 days of chronic sleep restriction." *Sleep.* 40.

¹⁴Bodenmann, S., Hohoff, C., Freitag, C., Deckert, J., Rétey, J.V., Bachmann, V., and Landolt, H.P. (2012). "Polymorphisms of ADORA2A modulate psychomotor vigilance and the effects of caffeine on neurobehavioral performance and sleep EEG after sleep deprivation." *British Journal of Pharmacology.* 165. Pg 1904-1913.

¹⁵Landolt, H.P., Rétey, J.V., and Adam, M. (2012). "Reduced neurobehavioral impairment from sleep deprivation in older adults: contribution of adenosinergic mechanisms." *Frontiers in Neurology.* 3. Pg 62-72.

¹⁶Porkka-Heiskanen, T., Strecker, R.E., Thakkar, M., Björkum, A.A., Greene, R.W., and McCarley, R.W. (1997). "Adenosine: a mediator of the sleep-inducing effects of prolonged wakefulness." *Science.* 276. Pg 1265-1268.

¹⁷Porkka-Heiskanen, T., and Kalinchuk, A.V. (2011). "Adenosine, energy metabolism and sleep homeostasis." *Sleep Medicine Reviews.* 15. Pg 123-35.

¹⁸Viola, A.U., Archer, S.N., James, L.M., Groeger, J.A., Lo, J.C., Skene, D.J., von Schantz, M., Dijk, D.J. (2007). "PER3 polymorphism predicts sleep structure and waking performance." *Current Biology.* 17.7. Pg 613.

¹⁹Rupp, T.L., Wesensten, N.J., Newman, R., Balkin, T.J. (2013). "PER3 and ADORA2A polymorphisms impact neurobehavioral performance during sleep restriction." *Journal of Sleep Research.* 22. Pg 160-165.

²⁰Mora-Rodriguez, R., Pallares, J.G., Lopez-Guillon, JM., Lopez-Samanes, A., Fernandez-Elias, VE., Ortega, JF. (2015). "Improvements on neuromuscular performance with caffeine ingestion depend on the time-of-day." *Journal of Science and Medicine in Sport.* 18. Pg 338-342.

²¹Garbaza, C., Benedetti, F. (2018). "Genetic factors affecting seasonality, mood, and the circadian clock." *Frontiers in Endocrinology (Lausanne).* 9.

²²Fuhr, U., Rost, KL., Engelhardt, R., Sachs, M., Liermann, D., Belloc, C., Janezic, S., Grant, D., Meyer, UA., Staib, AH. (1996). "Evolution of caffeine as a test drug for CYP1A2, NAT2, and CYP2E1 phenotyping in man by in vivo versus vitro correlations." *Pharmacogenetics.* 6. Pg 159-176.

²³Brager AJ, Ratcliffe RA, Grant D, Satterfield B, Capaldi V, Balkin T, Petrovick M. (2018). "2017 Genetic polymorphisms of sleep resiliency, sleep intensity, morning preference, and caffeine sensitivity are not associated with neurobehavioral performance under repeated cycles of total sleep deprivation." *Sleep.* 41. A7-A8.

Disclaimer

This summer fellowship was funded by the Army Outreach Education Program at the Walter Reed Army Institute of Research. The views expressed in this perspective piece solely represent those of the individuals and not the overall viewpoints of the Department of Defense and United States Army.

A Review of Treatments of Substance Use Disorder for Transitional Age Adults

BY KAI TREPKA, FATIMA FAIRFAX, JEFFREY HE, EUN JAE KIM, MINGU KIM, JEONGMIN LEE, RAJDEEP TRILOKEKAR, AND KATHERINE VENTURO-CONERLY

HARVARD UNIVERSITY

Abstract

Mental health concerns often develop during the “transitional age” (between 18-24 years old), a time during which young adults often begin to experience the stressors and responsibilities of emerging adulthood. Yet, treatment complications and limitations that uniquely affect this age group have frequently been unaddressed, both in academia and in practice. Substance Use Disorder (SUD) is one of the most common, costly, and intractable mental health concerns that affect this demographic. Here, we perform a literature review of 35 publications and find that traditional methods of treatments such as medications, family therapies, and Cognitive Behavioral Therapy (CBT) are not only underutilized, but also poorly applied within the transitional age demographic. Emerging treatment research suggests that devising treatment plans for transitional age youth using modified traditional practices and new evidence-based practices should improve the outcomes of SUD treatments. To improve treatment efficacy and adherence, we suggest promising areas of research surrounding therapeutic alliances, community engagement, continuity, Motivational Enhancement Therapy, and Multisystemic Therapy.

Introduction

Prevalence and severity of mental health concerns amongst adult and adolescent populations are well-established in research and practice. The distinctions between the two age groups has increasingly become a focus of mental health research, while mental health concerns are often overlooked in the gap period known as the “transitional age.” Here, we define the transitional age individuals to be those between 18-24 years of age, based on an analysis of when most mental health concerns first manifest and the consideration by the Centers for Disease Control and Prevention (CDC) and National Institute of Mental Health (NIMH) of an individual as a child or adolescent.¹ The CDC considers an individual to be an adolescent until they reach age 18.² The vast majority of mental health concerns develop between ages 15-24.^{1,3} Given these ranges, the intersection between childhood and mental illness’ nascent stages creates clear boundaries between adolescents, transitional age individuals, and fully grown adults. While research into adolescent mental health concerns is highly specific, transitional age individuals often lack a clear identity in the research lens and are grouped with one of the other two categories.

Longitudinal demographic studies from 1990 to 2010 have indicated that approximately one in five adolescent and transitional age individuals will experience mental illness.² Overall, these studies indicate

that the mental health burden has been increasing in developed countries, specifically upon these younger individuals. These individuals are at elevated risk for concerns such as depression, anxiety, substance abuse, ADHD, eating disorders, PTSD, etc. Often, the concerns increase in complexity as they have a high rate of comorbidity.⁴

One characteristic that influences the behavior of transitional age adults is an unstructured development period of individuation with fewer rules, which may lead to increased opportunity for substance use.⁴ Transitional age individuals also have fully developed limbic systems in the brain (which function for motivation and reward) accompanied by less developed prefrontal cortices (aiding in planning and inhibition), biologically predisposing them to take more risks with immediate reward without considering later consequences.⁵ They are also less likely to adhere to treatment plans and to view their recovery as possible and essential since they lack (a) the level of executive functioning capacity that a fully developed prefrontal cortex will later provide, (b) the motivation of retaining a marriage or a full-time career, or (c) the extrinsic motivation of parentally-imposed rules. This is highlighted by the highly stressful environments of college, the entry to adult life, and possible failure to live up to social expectations.⁴ Furthermore, transitional age individuals are generally not vocationally established, leading to a degree of continued dependence

on the systems of authority in place when they were minors. This makes treatment more complex, since it involves a greater circle of effected and influential individuals, but also opens new avenues of treatment through social systems.⁶ Finally, many individuals with substance use disorders (SUDs) use substances as coping mechanisms for general life stressors and therefore never develop other social and behavioral means of coping, which predisposes them to comorbid mental health disorders.⁷

There is a decline in utilization of mental health services as youth transition into young adulthood. For 16 and 17-year-olds, the yearly admission rate for inpatient, outpatient, and residential services in 2008 was 34 per 1000, while the number decreased dramatically to just 18 per 1000 for ages 18 and 19. Notably, among 20- and 21-year-olds who utilize mental health services, more were referred from the criminal justice system than from family or friends.⁸ This shows that there need to be more resources and support allocated to ensure that young adult clients can remain in treatment as they age out of child mental health systems.

Furthermore, there is a median delay of 11 years between the onset of a mental health concern and accessing services to correct it.³ Eisenberg et al. assert that most mental health disorders have their onset between ages 15-24, with transitional age individuals being uniquely vulnerable and important for targeted treatments. They

found that among college-attending students only approximately 36% of individuals who screened positive for depression or anxiety had accessed aid. This clearly indicates a gap that is wider among youth without access to the support structures postsecondary education provides.

Transitional age individuals, who experience a period of increased autonomy and decreased regulation, are the demographic in which substance misuse is uniquely prevalent. Yet, traditional treatments such as cognitive behavioral therapy, support groups, and contemporary medication prescription have been found to be often the least effective for these individuals.⁹ Thus, it is important to more completely investigate literature to identify gaps in current methods of treating substance misuse in transitional age adults. In this review, we investigate the shortcomings of current treatments for transitional age adults and identify promising new areas of research as a way forward.

Methods

Research papers on substance use disorders in transitional age adults were identified from an electronic database by using search terms to specify relevant articles. Using PubMed and Google Scholar (January 2000 to April 2017), combinations of the search terms transitional age, adolescent, substance use disorder, substance abuse, adolescent, and mental health were necessitated. For the purpose of this review, “transitional age” individuals are individuals between 18 and 24 years of age. Research papers focused on primarily non-transitional age individuals (i.e. individuals outside the age range of 18-24) were excluded. Of the remaining accessible articles, case studies and articles where mental health, non-transitional age individuals, or substance use were not central were excluded. Articles used in this review only had human subjects and were written in English.

Results

Traditional Methods

Because of the high risks of substance misuse and its corresponding impacts in young adults, improved SUD treatment for this age group is important. SUDs are not only most prevalent in the transitional age group, but also many treatments that are effective for adults and adolescents have been found to have severe shortcomings in treating this specific population.¹⁰

There are two prevalent pharmacological methods for the treatment of SUDs. As of 2009, most drugs to treat addiction are opioid agonists that eliminate withdrawal symptoms; the most commonly used drugs of this class are methadone and buprenorphine.¹¹ Post-treatment fidelity (continued sobriety) is the primary concern

of the effectiveness of pharmacological treatments. Naloxone, usually a drug prescribed to mitigate the effects of an overdose, can be used to overcome some of these problems, maintaining fidelity to treatment and preventing relapse.¹² However, Naloxone functions by inhibition of pattern completion responses which are crucial to learning, which indicates that it is not necessarily a good complement to a methadone regimen for transitional age individuals.¹² In other words, Naloxone is not the best pharmacological treatment for transitional age individuals; instead, physical withdrawal symptoms should be managed by methadone or buprenorphine as other routes of mental health treatment are pursued.

Many behavioral therapies exist independent of pharmacological approaches. Some evidence-based treatments include contingency management therapies, cognitive-behavior/skills-training therapies, motivational interviewing, and family treatments.⁹ Contingency management therapies motivate patients to perform desired actions, such as passing a drug test, by providing systemic rewards. Such a method has shown promise, but it ideally requires more outreach to policymakers and the general legal framework. Skills-training therapies have strong empirical backing in other mental health treatments, such as depression or anxiety, but there has been less study as to their limitations. Motivational interviewing, which entails guided conversation meant to increase intrinsic motivation and aid application of said motivation, has been shown to generally increase fidelity to treatments as it creates an overall more positive attitude towards treatment, thus staving off relapse, but has not been shown to be as effective on its own. Family treatments have shown greater efficacy on their own, but their main impact has also been shown to lie in fidelity, with the added effect of improving the mental health state of peripheral individuals impacted by the disorder.

Other studies have investigated the availability and use of mental health services for concurrent disorders amongst transitional age individuals.¹³ It was shown that 65% of the participants in their study with substance use disorders have at least one lifetime mental disorder. One such example where treatment of a concurrent mental health disorder affected SUD risk can be found in the complex relationship between medication for Attention Deficit Hyperactivity Disorder (ADHD) and the risk of substance misuse. Wilens et al. (2013) found in a meta-analysis of six studies that while pharmacological treatment decreased the risk in children with ADHD for later substance use disorder, this

protective effect decreased significantly as individuals transitioned from adolescence into adulthood.¹⁴ Mean onset of SUD in ADHD individuals occurs at 19 years old, which underlies the importance of factors surrounding transitional age (changing involvement of family, doctor, pharmacological, and other environmental supports) when considering complicating factors regarding SUD.

Despite being part of an age group with the highest level of need for mental health services, young adults are the least likely individuals to use mental health services. This demonstrates a need for integrated treatment models that address the intersection of substance use and mental illness. Overall, the rate of mental health treatment for transitional age individuals is low - only 15% of 18-25-year-olds receive treatment for depression, vs 38% of adolescents.¹⁵ Few studies of transitional age individuals have examined differing effects of mental health treatment between transitional age socioeconomic and ethnic subgroups. Although young adults have an overall higher prevalence of and lower treatment of alcohol and serious psychological distress, males (as opposed to females) and non-Whites (as opposed to Caucasians) both proportionally receive less treatment, even after accounting for economic and insurance disparities. This suggests that stigma and personal/cultural factors play a large role, potentially having negative effects on treatment fidelity. Surprisingly, income plays little role in the rate of treatment, perhaps because these individuals still have access to their parents' resources.

Race influences current treatment availability and fidelity. Employing a targeting strategy driven by race may be beneficial.¹⁶ Cost and stigma were the most significant barriers to receiving mental health care, and transitional age individuals are particularly susceptible to believing and being negatively affected by negative stigmas surrounding mental health. For example, an individual whose cultural background and home environment both stigmatize mental health disorders and/or addiction is less likely to seek treatment. Additionally, the study suggests that there exists a significant correlation between people who perceive that they need mental health treatment services but fail to receive them and people with substance use disorder (SUD). We believe that further studies are necessary to validate this claim.

Education also affects incidence and treatment. Education attainment and school enrollment are protective factors as substance misuse is less compatible with college student lifestyles.¹⁷ Treatment fidelity was higher

for college students. These students were more likely to complete substance misuse treatment than non-students and completed it in a faster timeframe. Students were most likely to misuse alcohol and marijuana as they can remain relatively high functioning under these substances, while non-students delved into other drugs.

It is unclear what effect regional differences have on substance misuse treatment. In literature prior to 2008, many studies showed that substance misuse was less common in rural areas in the US.^{18,19} However, a study from 2015 suggests that substance misuse is higher in rural areas, especially with alcohol and methamphetamine.¹⁹ In order to help youth in rural areas, intervention programs taking into consideration the rural context and the existing rural infrastructure need to be developed.

After the initial administration of treatment, close monitoring of potential for relapse is important. Two-thirds of the participants in the study had relapsed to drug use within six months.²⁰ While rates of relapse following treatment are comparably high in both adults and adolescents, the reasons for relapse differ significantly. Adults relapsed in social situations in which they experienced urges and temptations to drink/use or when they were trying to cope with a negative emotion and urges and temptations to drink/use. In contrast, nearly 70% of adolescent subjects reported that they relapsed in social situations when they were trying to enhance a positive emotional state.²¹ Given that young adults are often in environments where substance misuse is not only tolerated but even glorified, there need to be programs that rigorously follow up with young people recovering from substance misuse.

Finally, there are various public strategies that synthesize psychological and pharmacological approaches and are currently employed by substance misuse treatment services to help adolescents and their families: pairing families with professionals, organizing programs to gather families together, and enlisting policy makers to help make decisions for families through having access to experts.²² In other words, the first technique is an approach to help one particular family while the second groups families together to create a self-supporting community. The latter technique can have long-term effects in multiple families, which was the case in states such as Wisconsin where the Child and Youth Substance Abuse Subcommittee was installed to aid families in getting the right substance misuse treatment the victim needs. However, even with these different techniques being employed through more than nine states, the issues of substance misuse and recovery persist.

Modifications

Engagement between counselors and patients should be increased where feasible. Therapeutic alliances between counselors and young adult patients play an important role in dealing with substance misuse.²³ A therapeutic alliance exists when “the therapy dyad is engaged in collaborative, purposeful work”.²⁴ A therapeutic alliance emphasizes the participation of both counselors and their patients in creating goals and therapy tasks, fostering an emotional bond between patients and their therapists. The strength of the counselor patient alliance correlated significantly with positive treatment outcomes in patients. Patients who began treatment with higher levels of motivation and self-efficacy to abstain from substances showed stronger therapeutic alliances. Problems of mandated and resistant clients will still exist in many cases; however, a starting point of dyadic therapy (i.e. using the patient's goals to guide a treatment plan) will likely improve care.

Likewise, active participation of transitional age patients in making decisions on what medications to take leads to reduced symptoms, higher self-esteem, and improved fidelity.²⁵ Conversely, top-down prescriptions result in the overuse of psychotropic drugs, which have unique and negative side effects during the transitional age. The support of advocates such as mental health providers during meetings with psychiatrists can improve deliberation, as can psychiatrists who are available outside of office time to provide quick answers to concerns about dosage and about side effects.

Community engagement and external focus are beneficial in avoiding and alleviating substance use and other forms of mental illness.²⁶ Interestingly, Hispanic teens who feel more responsible for others (such as their family) are less likely to binge drink and smoke marijuana than teens who see emerging adulthood as a time for self-exploration or who feel less of an obligation to others.²⁶ Mentally ill teens also expressed much greater life satisfaction if they were engaged in work which they thought might benefit others and if they felt that they had meaningful social connections. This sense of belonging and usefulness is essential for mental health and healthy behaviors and can improve treatment effectiveness, though its effects may vary widely between different ethnic and social groups.^{27,28}

Continuity of treatment from adolescence into adulthood is important. More longitudinal, controlled health services research must be conducted to find optimal service models of treatment for those in the transitional age period.²⁹ More specifically, there needs to be systematic and seamless

transition protocols for patients with mental disorders who need to continue care into adult mental service. Currently, there is no consensus on what constitutes as a successful transition. Moreover, there is a decline in the use of mental health services by transitional age youth as they transition from adolescence to adulthood, even in the case of continuing disorders.

Varying the location of treatment alters its effectiveness, particularly for young adults receiving substance use treatment who also have co-occurring psychiatric disorders.⁹ People with co-occurring disorders (COD) were seen to put strain on coping resources, which resulted in poorer treatment response and outcomes. COD patients reported higher level of dependence severity and substance use consequences. For treatment, they reported that treatment delivered in familiar, residential settings could most benefit COD patients, as it alleviates some psychological stress by reducing environmental demand and providing a consistent emotional support system. Dedication to substance abstinence resulted in greater availability of cognitive resources and interventions were maximized by addressing motivation, self-efficacy and coping skills. The study was consistent with other articles showing that young adults are very susceptible to co-occurring disorders, and that patients with COD need specialized treatment for their health.

Increasing focus on Motivational Enhancement Therapy (MET) could be potentially effective for this age group. MET is a technique that relies on inducing rapid and internally motivated change, often causing individuals to overcome their opposition to engaging in treatments for substance misuse.³⁰ Tobacco addiction almost always begins between the ages of 18 and 24 and complicates mental health treatment by increasing the metabolism of antipsychotic and antidepressant medications. Of psychiatric patients, 50% with anxiety and 66% with depression smoke. Parents are a factor which enables smoking, both by purchasing cigarettes, providing a role model who smokes, and not discouraging tobacco use, especially within certain cultural groups. Additionally, although mental health providers and youth agree that smoking is a way for subjects to engage with their peers, it is also an addictive and unhealthy behavior, and providers and youths often propose different solutions to addiction. Youth often prefer to quit “cold turkey” and avoid pharmacologic treatments, whereas providers either view smoking as an inevitable “fact of life” or prioritize the use of cessation medication like nicotine patches. However, “motivational strategies” were recommended by both youth and providers, potentially

an avenue for compromise. Mental health facilities are also viewed by both providers and youth as a safe space to talk about these problems and are another potential way to address addiction.³¹

Similarly, there is often an escalation of drinking and alcohol use problems in the transitional age group.³² Despite high rates of alcohol usage, only a small proportion of people seek and receive treatment. Furthermore, canonical interventions fail for older adolescents. Late adolescence is a period of high vulnerability as there are vast “neurologic, cognitive, and social changes” that occur. Hence, treatment should capitalize on “transient high motivational states of youths” who, at a younger age, have fewer drug use problems that might arise from alcohol and are more likely to want to change their drinking habits. “Motivational enhancement techniques” (such as MET) should be considered as an early treatment.

Integrating multiple therapies is a potentially effective technique. For example, Hersh et al (2013) investigated a combination of Cognitive Behavioral Therapy (CBT) and MET.³³ Few studies have investigated how depression and conduct disorder affect efficacy of substance misuse treatment in transitional age individuals – prior research focused on depression and conduct disorders individually, instead of treating them as an integrated, unique problem. Cognitive behavioral therapy that addresses youths’ hostile attribution bias and builds social support combined with motivational enhancement therapy to increase goal setting proved effective, even with only four sessions over a year. However, this study lacked a control group and focused mostly on Caucasians. Interestingly, contrary to previous studies, stronger depression makes substance misuse treatment more effective, implying both that more research needs to be done on the way mental health disorders can affect substance misuse treatment and that one must consider the length of treatment and follow up when comparing different studies. Higher income is also correlated with better treatment outcomes, possibly because wealthier youths have more opportunities to remove themselves from negative peer associations.

Multisystemic Therapy (MST), a specific family-involved therapy, has some significant benefits down the line, but is not universally helpful on its own.³⁴ Firstly, it indicates that there is a disaggregation of antisocial tendencies where the problems associated with the drug misuse, namely crime, are transitioned from an “overt” setting to a “covert” setting. That is to say, the patients who go through MST are less likely to commit aggressive crime, but no less

likely to commit property crimes. This, along with a lack of significant results in terms of improving psychiatric well-being, indicates that MST is likely to be only useful as a complement as long as the underlying issues are treated.

To further accentuate this nuanced point of view, though, the efficacy of MST is also sensitive to cultural needs and is suitable for a diverse range of ethnic groups.³⁵ This complements the fact that in the previous study 60% of the patients were African American. The only major racial factor that had an impact was that race matching between therapist and patient was beneficial. Rowe et al’s work validated the persistent efficacy of MST for up to 14 years.³⁵

This is because, most importantly, MST’s effects are persistent over many years. Thus, combining MST with other treatment forms can ensure fidelity to treatment and preventing relapses.¹² The patients in the Henggeler study had low treatment fidelity to the MST, but that could be explained by the lack of treatment for the underlying factors.³⁵ Thus, MST is undeniably beneficial, but only truly makes an impact in the context of a larger treatment plan.

Discussion and Conclusions

In this paper, we report that the transitional age demographic is socially and physiologically distinct from both adolescents and adults. Usually, when a large age demographic with a disorder exists, an outline for treatment rapidly emerges, as it did for adults and adolescents with mental health disorders. Despite the large size of the transitional age demographic and the propensity of mental health and substance misuse disorders to arise during this time, there is still no unified framework to treat transitional age individuals. Treatments for co-occurring substance misuse and mental health disorders need to account for their uniqueness.

To prevent disorders, therapists should screen clients for different SUDs since young adults don’t always recognize the importance of treating their disorders early.³² In treatment, therapists should adopt a fluid, patient-centered approach, use developmentally tailored communication and engagement, involve and train parents if possible, and help patients to develop a non-using support network. Finally, therapists should emphasize treatment engagement rather than full adherence, as young adults often relapse at first and cannot immediately reach recovery. Successful recovery in transitional age individuals should not be defined as immediate, uninterrupted sobriety, but instead as a continued dedication to treatment and desire to become sober, healthy, and productive. This provides

a strong metric by which we can evaluate current and potential treatment avenues.

Traditional methods such as pharmacological regimes and late-stage contingency management are relatively ineffective in terms of treatment fidelity. Family treatments, motivational interviewing, and CBT are underutilized and do not account for relapse. As a result, current treatment overall is underutilized and results in racial and socioeconomic discrepancies.

Instead, treatments should be modified to focus more on creating therapeutic alliances, engaging patients with the community and youth, continuing treatment from youth to adulthood, deciding the most appropriate physical location of treatment, increasing focus on MET and CBT (perhaps as an integrated therapy), and using multisystemic therapy to supplement other treatments.

Overall, we found a lack of research on treatment specific to the period of transition from youth to adulthood. Where research was present, it was often solely correlational, and the sample size rarely exceeded a few dozen people. Even though the intent of the research was to improve treatment and clinical outcomes, it often lacked essential features to verify its efficacy, such as clinical trials with a control group. Finally, the methods that were proposed are not implemented widely by psychiatrists today. More clinical research is needed on the transitional age demographic to understand why certain treatments are ineffective and which modifications should be made.

An essential limitation of our review paper is that in our discussion of transitional age individuals, we lack substantial overview of the effects of childhood and adolescence on the lives, habits, and personalities of those in the transitional age. As a result, we cannot adequately capture the potential importance of mitigating risk factors for SUD that present themselves in childhood and adolescence. If we take this logic to its conclusion, substantive treatment for transitional age individuals is potentially of a lower priority as deeper knowledge and preventative treatment rooted in childhood and adolescence could prevent both SUD and associated concerns.

However, this logic is predicated upon further research in transitional ages that is not currently showcased in literature. This reinforces the importance of our thesis calling for further research. Further research, which our paper could not capture, could encapsulate dichotomies between university-attending individuals, vocational individuals, or unemployed individuals. It could also include longitudinal studies and treatment pipelines which have yet to be

publicized (clinical/lab phases) and were thus not expressed in this paper. To address this concern, we searched the intellectual property database Innography twice, once for “substance use mental health” and once for “substance use disorders addiction treatment” (Figure 1). Both showed similar trends in both patents filed and grants given; the latter showed significantly greater amounts of data and is displayed below. The number of grants offered spiked in 2012 and has steadily declined since. Overall, filings follow grants with a one- to two-year delay, and the four-year decline since 2012 in both grants and patents indicates that the field is “drying up” and needs revitalization. This positions research into transitional age individuals as a uniquely important avenue of research. Findings in this specific avenue of research could stimulate the idle academic status quo. Given the established lack of substantial information, which led to the limitations we highlighted in our review, a stimulus could lead to new developments in addiction and mental health treatment beyond the narrow scope of transitional age individuals and shed new light on our understanding of comprehensive prevention and treatment methodologies.

Further, our analysis of the primary studies in that the range of age group identified as transitional youth varied between studies. The conclusions we have drawn above stem from a small available selection of studies. There is a need for more research on transitional age individuals and their relationship with substance use in general, and we hope that currently ongoing longitudinal studies can shed new light on

this critical issue.

While it is outside the scope of our paper, it may be important as well to consider experiences and treatments provided during childhood and adolescence for these transitional age individuals, as problems during transitional age may more commonly stem from previous experiences. This, if found to be true, would suggest that efforts focused on preventative care during earlier years would be better served than those focused on treatments targeting transitional age individuals.

Even after this review of multiple studies on substance misuse of individuals who are in the transitional age group, there are still many questions left unanswered. A major issue includes the multitude of factors that may take role in an adolescent’s decision. Transitional age has proven to be a particular age group that can be found to be committing detrimental behaviors such as substance misuse. However, the elements in the adolescent’s environment that can elicit such behaviors are yet to be determined, whether these elements depend on ethnic culture, gender, geography, or race just to name a few. Additionally, part of the difficulty individuals of transitional age encounter in coping with issues such as substance misuse include their lack of utilization of existing mental health services. What specifically yields these decisions have not been thoroughly documented yet. Finally, while there are many similarities in the social and molecular mechanisms contributing to substance misuse, important differences in the research, treatment, and legal environments exist depending on the substance being misused

(e.g. alcohol vs marijuana). While this review primarily focuses on unifying factors that impact both alcohol and marijuana use disorders in transitional age adults, research and treatment must also take into account the particular substance being abused.

Thus, we recommend future research should be conducted with longitudinal studies following a large population of youth that finds correlations between factors in childhood and transitional age substance misuse disorders. A correlational study with a large sample size would greatly add to our understanding of the cause for these disorders, and a greater emphasis on prevention can be made, rather than treatment. In addition, transitional age individuals are admitted for mental health treatment at extremely lower rates; more research is needed to explain these low admission rates. More information addressing the factors of deterrence that most transitional age individuals face when considering mental health treatments can help health care workers propose more effective plans that would benefit all potential treatment options. Finally, clinical trials with the treatments listed in our results section would allow a robust evaluation of each treatment option.

References

- ¹[N.A.] (2020) “Child and adolescent mental health.” National Institute of Mental Health. National Institutes of Health. <<https://www.nimh.nih.gov/health/topics/child-and-adolescent-mental-health/index.shtml>> (Accessed 02/05/2020)
- ²Perou, R., Bitsko, R. H., Blumberg, S. J., Pastor, P., Ghandour, R. M., Gfroerer, J. C., Hedden, S.L., Crosby, A.E., Visser, S.N., Schieve, L.A., Parks, S.E., Hall, J.E., Brody, D., Simile, C.M., Thompson, W.W., Baio, J., Avenevoli, S., Kogan, M.D. and Huang, L.N. (2013) “Mental health surveillance among children—United States, 2005–2011.” *Morbidity and Mortality Weekly Report* 62, Pg 1-35.
- ³Eisenberg, D., Golberstein, E. and Gollust, S. (2007) “Help-seeking and access to mental health care in a university student population.” *Medical Care* 45(7). Pg 594-601.
- ⁴Mandarino, K. (2014). “Transitional-age youths: Barriers to accessing adult mental health services and the changing definition of adolescence.” *Journal of Human Behavior in the Social Environment* 24(4). Pg 462-474.
- ⁵Chung, W. W. and Hudziak, J. J. (2017) “The transitional age brain: the best of times and the worst of times.” *Child and Adolescent Psychiatric Clinics* 26(2). Pg 157-175.
- ⁶Kim, H., Munson, M. R. and McKay, M. M. (2012) “Engagement in mental health treatment among adolescents and young adults: a systematic review.” *Child and Adolescent Social Work Journal* 29(3). Pg 241-266.
- ⁷Hartz, S. M., Pato, C. N., Medeiros, H., Cavazos-Rehg, P., Sobell, J. L., Knowles, J. A., Bierut, L.J. and Pato, M. T. (2014) “Comorbidity of severe psychotic disorders with measures of substance use.” *Journal of the American Medical Association Psychiatry* 71(3). Pg 248-254.
- ⁸Pottick, K.J., Bilder, S., Vander Stoep, A. et al. (2007) “US patterns of mental health service utilization for transition-age youth and young adults.” *The Journal of Behavioral Health Services and Research* 35. Pg 373-389.
- ⁹Carroll, K. M. and Onken, L. S. (2005)

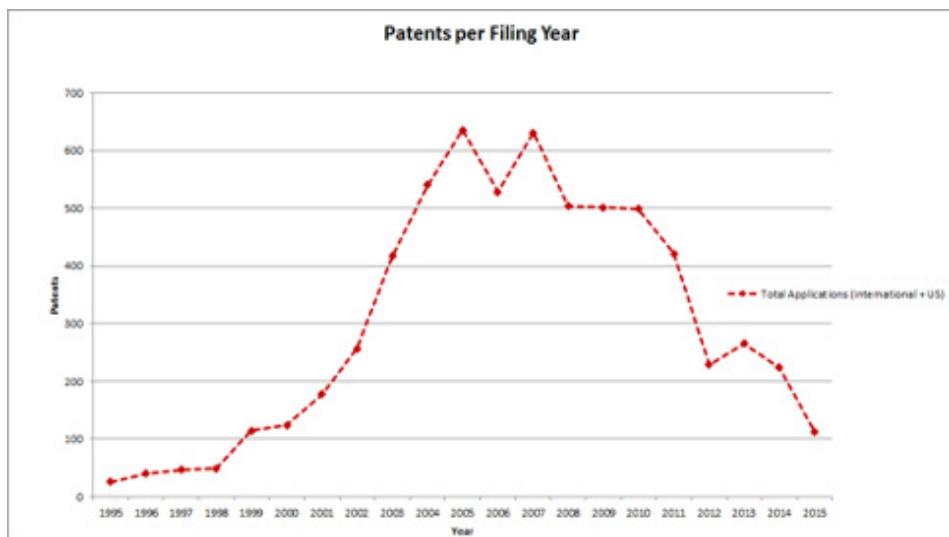


Figure 1: Substance Use Disorder Patent Applications. Global patent applications for “substance use disorders addiction treatment” from 1995 to 2015. Since 2012, there has been a steady downward trend, suggesting a need for revitalization of funding SUD research.

“Behavioral therapies for drug abuse.” *American Journal of Psychiatry* 162(8). Pg 1452-1460.

¹⁰Bergman, B. G., Greene, M. C., Slaymaker, V., Hoeppner, B. B. and Kelly, J. F. (2014) “Young adults with co-occurring disorders: substance use disorder treatment response and outcomes.” *Journal of Substance Abuse Treatment*, 46(4). Pg 420-428.

¹¹[N.A.] (2009) “Guidelines for the psychosocially assisted pharmacological treatment of opioid dependence.” *World Health Organization*. https://www.who.int/substance_abuse/publications/9789241547543/en/ (Accessed 02/05/2020)

¹²Kesner, R. P., Kirk, R. A., Clark, J. K., Moore, A. and Keefe, K. (2016) “Naloxone injections into CA3 disrupt pattern completion associated with relapse from cocaine seeking.” *Hippocampus* 26(7). Pg 892-898.

¹³Miller, S., Ringeisen, H., Munoz, B., Hedden, S., Colpe, L., Rohloff H. and Embry, V. (2016) “Correlates of mental health services use among young adults with mental illness: results from the National Survey on Drug Use and Health.” *Psychiatric Services*, 67(6). Pg 642-649.

¹⁴Wilens, T. E., Faraone, S. V., Biederman, J. and Gunawardene, S. (2003) “Does stimulant therapy of attention-deficit/hyperactivity disorder beget later substance abuse? A meta-analytic review of the literature.” *Pediatrics* 111(1). Pg 179-185.

¹⁵Adams, S., Knopf, D. and Park, M. J. (2014) “Prevalence and treatment of mental health and substance use problems in the early emerging adult years in the United States: Findings from the 2010 National Survey on Drug Use and Health.” *Emerging Adulthood* 2(3). Pg 163-172.

¹⁶Mason, M.J., Keysey-Marcus, L., Snipes, D., Benotsch, E. and Sood, B. (2013) “Perceived mental health treatment need and substance use correlates among young adults.” *Psychiatric Services* 64(9). Pg 871-877.

¹⁷Sahker, E., Acion, L. and Arndt, S. (2015) “National analysis of differences among substance abuse treatment outcomes: College student and nonstudent emerging adults.” *Journal of American College Health* 63(2). Pg 118-124.

¹⁸Borders, T. F. and Booth, B. M. (2007). “Rural, suburban, and urban variations in alcohol consumption in the United States: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions.” *The Journal of Rural Health* 23(4). Pg 314-321.

¹⁹Lambert, D., Gale, J. A. and Hartley, D. (2008) “Substance abuse by youth and young adults in rural America.” *The Journal of Rural Health* 24(3). Pg 221-228.

²⁰Cornelius J.R., Maisto S.A., Pollock N.K., Martin C.S., Salloum I.M. and Lynch K.G. (2001) “Rapid relapse generally follows treatment for substance use disorders among adolescents.” *Addictive Behaviors* 27. Pg 1-6.

²¹Ramo, D. E. and Brown, S. A. (2008) “Classes of substance abuse relapse situations: A comparison of adolescents and adults.” *Psychology of Addictive Behaviors* 22(3). Pg 372-379.

²²McGillicuddy, and Eliseo-Arras. (2012) “Parent-adolescent report correspondence on adolescent substance abuse among teens in residential rehabilitation.” *Addictive Behaviors* 37(4). Pg 456-462.

²³Urbanoski, K. A., Kelly, J. F., Hoeppner, B. B. and Slaymaker, V. (2012) “The role of therapeutic alliance in substance use disorder treatment for young adults.” *Journal of Substance Abuse Treatment* 43(3). Pg 344-351.

²⁴Hatcher, R. L. and Barends, A. W. (2006) “How a return to theory could help alliance research.” *Psychotherapy: Theory, research, practice, training* 43(3). 292-299.

²⁵Delman, J., Clark, J., Eisen, S. and Parker, V. (2015) “Facilitators and barriers to the active participation of clients with serious mental illnesses in medication decision making: The perceptions of young adult clients.” *Journal of Behavioral Health Services and Research* 42(2). Pg 238-253.

²⁶Allem, J. P., Lisha, N. E., Soto, D. W., Baezconde-Garbanati, L. and Unger, J. B. (2013) “Emerging adulthood themes, role transitions and substance use among Hispanics in Southern California.” *Addictive behaviors* 38(12). Pg 2797-2800.

²⁷Resnicow, K., Soler, R., Braithwaite, R. L., Ahluwalia, J. S. and Butler, J. (2000). “Cultural sensitivity in substance use prevention.” *Journal of community psychology* 28(3). Pg 271-290.

²⁸Unger, J. B. (2015). “Preventing substance use and misuse among racial and ethnic minority adolescents: Why are we not addressing discrimination in prevention programs?” *Substance Use and Misuse* 50(8-9). Pg 952-955.

²⁹Reale, L. and Bonati, M. (2015). “Mental disorders and transition to adult mental health services: A scoping review.” *European Psychiatry* 30(8). Pg 932-942.

³⁰Tevyaw, T. O. L. and Monti, P. M. (2004). “Motivational enhancement and other brief interventions for adolescent substance abuse: foundations, applications and evaluations.” *Addiction* 99. Pg 63-75.

³¹Prochaska, J., Fromont, S., Wa, C., Matlow, R., Ramo, D. and Hall, S. (2013) “Tobacco use and its treatment among young people in mental health settings: A qualitative analysis.” *Nicotine and Tobacco Research* 15(8). Pg 1427-1435.

³²Brown, S. A., McGue, M., Maggs, J., Schulenberg, J., Hingson, R., Swartzwelder, S., Martin, C., Chung, T., Tapert, S.F., Sher, K., Winters, K.C., Lowman, C., and Murphy, S. (2008) “A developmental perspective on alcohol and youths 16 to 20 years of age.” *Pediatrics* 121. Pg S290-S310.

³³Hersh, J., Curry, J. and Becker, S. (2013) “The influence of comorbid depression and conduct disorder on MET/CBT treatment outcome for adolescent substance use disorder.” *International Journal of Cognitive Therapy* 6(4). Pg 325-341.

³⁴Henggeler, S. W., Clingempeel, W. G., Brondino, M. J. and Pickrel, S. G. (2002) “Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders.” *Journal of the American Academy of Child and Adolescent Psychiatry* 41(7). Pg 868-874.

³⁵Rowe, C. L. (2012) “Family therapy for drug abuse: Review and updates 2003-2010.” *Journal of Marital and Family Therapy* 38(1). Pg 59-81.

Understanding Benthic Macroinvertebrate Communities in Beaver Meadows of Rocky Mountain National Park

BY VALERIE T. DOEBLEY
COLORADO STATE UNIVERSITY

Abstract

In mountain regions, river-floodplain meadows can develop in unconfined valley sections of the river network. River-floodplain meadows are ecologically significant regions due to their ability to retain carbon and nutrients through several processes, including the storage of large woody debris and the accumulation of particulate and dissolved organic carbon in the sediment during flood events. These processes are influenced by the historic land use, such as livestock grazing or farming, and the presence of beavers in the meadow. Past projects in Rocky Mountain National Park examined the physiochemical and hydrologic characteristics of river-floodplain meadows with current or past beaver activity. However, past studies did not look at the communities of aquatic organisms residing in these areas.^{1,2} Aquatic macroinvertebrates are an effective biological indicator that are used to examine the effects of anthropogenic impacts on lotic ecosystems. The purpose of this study is to examine the aquatic macroinvertebrate communities at the inflows and outflow of six beaver meadows of Rocky Mountain National Park. By studying the aquatic insect communities at both ends of active and abandoned beaver meadows, this project examines: 1.) How the location above or below a meadow impacts communities, by comparing community composition at the outflows to inflows of meadows; 2.) any non-location-specific impact of beaver activity by comparing active and abandoned meadow communities, and, 3.) the interaction between location and beaver activity, in other words, a comparison of the change between inflow and outflow communities in active versus abandoned meadows. There were significant differences in several taxa between inflows and outflows of meadows, suggesting that meadows do impact the downstream communities. There were also more families that displayed overall higher abundances in active than abandoned meadows, regardless of inflow or outflow location. There were also two families, Glossosomatidae and Leptophlebiidae, which displayed an interaction effect such that their abundance significantly increased between the inflow and outflow of active meadows, but not of abandoned meadows.

Introduction

River-floodplain meadows in unconfined valley sections of mountain stream systems are an important sink in the carbon cycle.¹ These areas are often impacted by human land-use practices, which can alter the natural dynamics of the river and meadow. The presence or absence of beavers is an important characteristic of river-floodplain meadows that is impacted by past and current land-use practices. Beaver activity is associated with increased geomorphic complexity, carbon retention, and aquatic ecosystem metabolism.² Another study found that the river-floodplain meadows with beaver activity are particularly effective at retaining and processing carbon due to several natural processes occurring in these meadows. These processes include the storage of large woody debris and the capture of dissolved and particulate organic carbon into sediments during flooding in unconfined sections.¹ The increased geomorphic complexity, aquatic ecosystem metabolism, and carbon retention in beaver meadows may impact downstream habitat and community composition of biological indicator organisms

such as benthic macroinvertebrates.

Examining the macroinvertebrate communities of streams is commonly used to biomonitor how humans impact stream health and water quality.³ A healthy stream is capable of supporting a diverse macroinvertebrate community with a variety of life history traits and feeding habits. Accordingly, abundance and diversity of macroinvertebrate taxa are biological indicators of stream ecosystem health. Comparing the macroinvertebrate communities of beaver-less meadows with more intense land-use histories to less impacted active beaver meadows can indicate how human activities affect stream quality in meadows.

Furthermore, a stream's characteristics impact the macroinvertebrate community composition due to the various habitat needs and trophic relations of different taxa.⁴ The functional feeding groups of the taxa found in a stream can provide insight to the stream's ecosystem and water quality. Since feeding groups such as shredders, collector-filterers, and collector-gatherers rely on coarse particulate organic matter (CPOM)

and fine particulate organic matter (FPOM) as food, it is possible that a meadow's effect on the particulate organic carbon in a stream is a mechanism impacting downstream macroinvertebrate community composition.⁵

The purpose of this study is to use benthic macroinvertebrate communities as a lens to examine the differences between the quality of streams at the inflows and outflows of six river-floodplain meadows in Rocky Mountain National Park (RMNP). Based on past studies on the physiochemical and hydrologic properties of river-floodplain meadows^{1,2}, I hypothesized that macroinvertebrate abundance and diversity would be greatest at the outflows of active beaver meadows. The six meadows in this study were selected because they have different land-use histories and consequently they have different levels of beaver activity and geomorphic complexity. These meadows were also selected because previous studies through Colorado State University's Natural Resource Ecology Laboratory already examined their physiochemical and hydrologic characteristics.^{1,2} By examining these meadows at both inflows and outflows,

it can be determined if processes occurring within these meadows are impacting the downstream macroinvertebrate communities. Furthermore, comparing the differences between inflows and outflows across active and abandoned beaver meadows can provide insight to the effects of current beaver activity and channel complexity on ecosystem and habitat quality.

Methods

A conceptual diagram outlining the research process for this study is in **Table 1**.

Study Location

Benthic macroinvertebrate sampling was performed at two places on six meadows in Rocky Mountain National Park, for a total of 12 sampling locations. These meadows were divided by activity level into two categories: active meadows and abandoned meadows. Active meadows have current beaver activity and included Mill Creek and Glacier Creek. Abandoned meadows no longer have any beaver activity and included Moraine Park, Upper Beaver Meadows, Hidden Valley and Cow Creek. All of the abandoned meadows had historic beaver activity and still share varying characteristics of active meadows, such as slowed movement of water.

Each meadow was sampled at two locations: an inflow location upstream of the meadow and an outflow location downstream from the meadow (**Figure 1**). Each meadow has one inflow and one outflow. Active meadow sampling locations are referred to as “active inflow” and “active outflow”. Abandoned meadow sampling locations are referred to as “abandoned inflow” and “abandoned outflow”. Classifications of the six meadows are shown in **Table 2**. Aquatic macroinvertebrates were sampled at both inflow and outflow of each meadow to help determine if processes occurring within a meadow result in differences between the macroinvertebrate communities upstream and downstream of the meadow. The final study design consisted of two active and four abandoned meadows, each with an inflow and outflow sampling location (**Figure 2**).

Field Methods

Benthic macroinvertebrate sampling was performed using a 500-micron Surber sampler and a 600-micron sieve. At each sampling location, Surber samples were collected in riffles or runs no deeper than knee height. Four replicates were taken at four suitable riffles or runs positioned closest to the stream gauging station at each location. Replicates were taken moving from downstream to upstream to avoid disturbing subsequent replicates. At each selected riffle or run: 1) the Surber sampler was positioned, 2) each rock and piece of cobble in the square plot was picked up and vigorously scrubbed by hand

so that the contents of the plot flowed into the net, and 3) any remaining sand, gravel, plants, or small rocks within the plot were stirred by hand.

After the entire contents of the plot were thoroughly scrubbed and stirred up, the Surber sampler was removed from the water. The sample was deposited into a five-gallon bucket of water by turning the net of the Surber sampler inside out and dipping it into the bucket. The contents of the bucket were then poured through a 600-micron sieve and transferred into a 1 L plastic jar. The Surber sampler, bucket, and sieve were inspected for any remaining macroinvertebrates, which were transferred into the jar using forceps. Samples were preserved in a 95% ethanol solution. Although lower concentrations of ethanol are adequate to preserve samples, 95% ethanol is used in the field because a small amount of water from the stream ends up the jar, diluting the ethanol's concentration. Once in the laboratory, individual specimens from the sample can be safely transferred into new vials containing an 80% ethanol solution.

Laboratory Methods

In the laboratory, all macroinvertebrates were removed from the sample jars containing 95% ethanol, sorted, and placed into new, separate vials containing 80% ethanol. Other organic materials from the sample were also preserved. All macroinvertebrates from the samples were sorted by taxa using a dissection microscope. Aquatic insect larvae from the samples were sorted by family. Pupas found in the samples were sorted by order. Some samples also included some organisms other than aquatic insects. These included Bivalves, Trombidiformes (water mites), Gastropods, Annelids, Collembola, and terrestrial insects.

Data Analysis

Data analysis was performed using R-studio. A series of one-way and two-way ANOVAs were run to compare the abundance of different taxa across different site categories: active vs abandoned meadows, inflow vs outflow locations, or an interaction (activity level x location), using an alpha-value of 0.1. P-values of less than 0.1 indicate that the difference in abundance between site categories is significantly significant. A p-value greater than 0.1 indicates that there is no significant difference in abundance between site categories. No significant difference could mean: 1.) the site category does not impact the abundance, or 2.) there was not enough data collected to find an impact on abundance statistically significant. For this study, four replicate samples were taken at 12 sites, categorized into two activity levels and two locations. While a greater number of replicates and sites might lead to more statistically significant results, it is still possible to achieve statistically significant results with a relatively

small number of sites, especially in cases where there is a strong and consistent relationship between abundance and site category.

For these analyses, the insects were grouped by order, family, or functional feeding group (**Table 3**). Functional feeding groups were determined using information provided in “The Aquatic Insects of North America”.⁶ All abundance values were logarithmically transformed by taking the natural log of the abundance plus one. Logarithmic transformations are a common way to account for the issue that benthic macroinvertebrate datasets are usually not normally distributed and do not meet the assumptions of parametric statistical methods (e.g., homogeneity of variance).⁷

Results

Total macroinvertebrate abundances were variable between locations, meadows, and in some cases between replicates, with densities ranging from 12 to over 600 specimens per Surber sample. The average abundance of macroinvertebrates per Surber sample in active meadows was 94 specimens at inflow sites and 197 specimens at outflow sites. In abandoned meadows on average there was 111 specimens at inflows and 110 specimens at outflows (**Figure 3**). The number of families of aquatic insects per sample was also examined. Active meadows averaged 9.0 families at inflows and 9.9 families at outflows. Abandoned meadows averaged 8.9 families at inflows and 9.1 families at outflows (**Figure 4**). Neither total abundance nor number of families per Surber sample was found to be statistically significantly different between inflows and outflows or across activity levels.

Macroinvertebrate samples contained aquatic insect larvae from five orders: Ephemeroptera (mayflies), Plecoptera (stoneflies), Trichoptera (caddisflies), Diptera (true flies), and Coleoptera (beetles). These orders are divided into families. The raw data by orders and families can be found in table 4. Three of the aquatic insect orders, Plecoptera, Ephemeroptera, and Coleoptera, were overall more abundant in active meadows than abandoned meadows, regardless of the location on the meadow (p-values = 0.063, 0.0608, and 0.00364, respectively). No orders showed significant differences between inflow and outflow abundances or an interaction effect between activity level and location. Although most of the organisms collected were aquatic insects, some samples also contained Bivalves, Trombidiformes (water mites), Gastropods, Annelids, Collembola, and terrestrial insects.

The aquatic insect larvae found in these samples belonged to several functional feeding groups including collector-gatherers, collector-filterers, shredders, scrapers, herbivorous piercers, and predators (**Figure 5**). In meadows

of both activity levels (active and abandoned) and at both locations (inflow and outflow) the most abundant feeding group was collector-gatherers, which accounted for more than half of the total insects collected during this study. The feeding habits of individual families can be found in **Table 3**, as well as the ANOVA results regarding each family.

Collector-Filterers

One noteworthy feeding group was the collector-filterers. For both active and abandoned meadows, the average number of collector-filterers was greater at the meadow's outflow than the inflow ($p\text{-value} = 0.0731$). Families of collector-filterers were Simuliidae, Hydropsychidae, and Brachycentridae. The dipteran family Simuliidae, which accounted for the majority of the collector-filterers, displayed the same pattern as the overall collector-filterer group between the inflows and outflows ($p\text{-value} = 0.0447$). The abundance of Simuliidae likely influenced the overall pattern for its entire feeding group.

The other two families of collector-filterers, Brachycentridae and Hydropsychidae, are both caddisflies and were less abundant than Simuliidae. Neither Brachycentridae nor Hydropsychidae were found to have significant differences in inflow and outflow abundance. However, Brachycentridae was significantly more abundant overall in active meadows than in abandoned meadows, when inflow and outflow abundances are averaged for each meadow. ($p\text{-value} = 0.000436$). The opposite was true of the family Hydropsychidae, which was significantly more abundant in abandoned meadows than in active meadows, regardless of the location on the meadow ($p\text{-value} = 0.0863$).

Collector-gatherers

The collector-gatherer feeding group contained the majority of aquatic insects in this study. While the collector-gatherer group overall did not have any differences between inflows and outflows or between activity levels, individual families within the group displayed some significant results. A collector-gatherer mayfly, Leptophlebiidae, had a significant interaction between location and activity level ($p\text{-value} = 0.0680$), in which the abundance was found to increase between inflow and outflow in the active meadows, and decrease between inflow and outflow in the abandoned meadows. Leptophlebiidae was also more abundant at the inflows than the outflows of meadows of all activity levels ($p\text{-value} = 0.0774$). Two other collector-gatherer families were significantly more abundant in active meadows than abandoned meadows, regardless of inflow or outflow location; these families were Baetidae ($p\text{-value} = 0.00757$), and Elmidae ($p\text{-value} = 0.0034$).

Piercers

The feeding group piercers were found

almost exclusively at abandoned meadows. The caddisfly Hydroptilidae was the only piercer family present in this study. This family was found to be significantly more abundant at abandoned meadows than active meadows ($p\text{-value} = 0.0539$).

Predators

While the predator feeding group overall did not produce any significant results, individual families within the predator group displayed some significant results. One such predator was the stonefly Chloroperlidae, which was more abundant in the active meadows than the abandoned meadows, regardless of inflow or outflow locations ($p\text{-value} = 0.0477$). Another predacious stonefly, Perlidae, displayed an interaction effect between activity level and location in which their abundance was significantly higher at the active inflow than any other site ($p\text{-value} = 0.0423$). A predacious dipteran, Dolichopodidae, displayed the same interaction behavior as Perlidae ($p\text{-value} = 0.0126$). Additionally, Dolichopodidae was significantly more abundant at inflows than outflows, regardless of the meadow's activity level ($p\text{-value} = 0.0725$). Although these results are statistically significant, it's important to consider that only six total Dolichopodidae specimens were found during this study, and all six were found at the inflow to Mill Creek, an active beaver meadow.

Scrapers

The scraper feeding group also displayed several interesting differences in distribution across the meadows. For both active and abandoned meadows, the average number of scrapers was greater at the meadow inflow than the meadow outflow ($p\text{-value} = 0.0008$). The families of scrapers included Heptageniidae, Ameletidae, Uenoidae, and Glossosomatidae. The scraper mayfly Heptageniidae was found to be significantly more abundant at the inflows than the outflows for meadows of both activity levels ($p\text{-value} = 1.29E-05$). Because Heptageniidae was a very abundant scraper family, Heptageniidae influenced the overall patterns seen in the distribution of scraper feeders.

In addition to Heptageniidae, there were several less abundant scraper families. Another scraper mayfly family, Ameletidae, was significantly more abundant in active meadows than abandoned meadows ($p\text{-value} = 0.0301$). It is worth noting that Ameletidae was very uncommon in this study and was only present at one active meadow, Hidden Valley, and one abandoned meadow, Cow Creek. The family Glossosomatidae is a scraper caddisfly which had a significant interaction between location and activity level ($p\text{-value} = 0.0691$), in which the abundance was found to increase between inflow and outflow in the active meadows, and decrease between inflow and outflow in

the abandoned meadows. Another scraper caddisfly family, Uenoidae, was only found three times during this study, twice at the inflow of an active meadow, Mill Creek, and once at the outflow of an abandoned meadow, Hidden Valley.

Shredders

The only shredder feeder that produced significant results was the stonefly Nemouridae. Nemouridae was significantly more abundant overall in active meadows than in abandoned meadows ($p\text{-value} = 0.0569$). While not statistically significant, Nemouridae also behaves as though it may have an interaction between inflow/outflow and activity level ($p\text{-value} = 0.1243$). Nemouridae was the most abundant shredder family from this study and influenced the overall patterns seen in the distribution of shredder feeders. Large amounts of variation between replicates caused the interaction for Nemouridae to not be statistically significant, even though in figure 5 this interaction appears quite drastic for shredders. The collector-gatherer Elmidae also showed the same insignificant pattern as Nemouridae, in which abundance appears to increase more between the inflow and outflow of the active meadows than the abandoned meadows. Although not always significant, the presence of this same pattern across several taxa should be noted due to the possibility of it not being due to random chance.

Discussion

The results of this study can be grouped into three categories regarding the macroinvertebrate community composition: differences across inflows and outflows, differences across active and abandoned meadows, and an interaction effect of these two factors. First, differences in the abundances of several taxa across inflows and outflows, regardless of meadow type, suggest that processes occurring within floodplain meadows are altering the downstream aquatic insect communities. Second, some taxa are more abundant in active meadows overall, regardless of inflow or outflow location, suggesting that some characteristic of study areas are affecting the streams upstream of the meadows as well as downstream of the meadows. Finally, for some taxa an interaction occurs in which active meadows have different changes between the inflow and outflow communities than abandoned meadows. This interaction suggests that some aquatic insect taxa are more sensitive to differences between active and abandoned meadows and the processes occurring within them that impact outflow habitat. The mechanisms causing these three different categories of results are not directly examined in this study. Information from past publications on the feeding habits and life history traits

of different aquatic insects can help predict possible mechanisms driving these differences in aquatic insect communities.^{8,9}

Differences between inflows and outflows

For some families, their average abundance was significantly different between inflows and outflows. The families in **Figures 6 and 7** demonstrate examples of this type of relationship. Due to their location, outflow habitats are subject to the ecological impacts of processes occurring in the meadows upstream, while inflow habitats are not. Meadow processes include storage and gradual breakdown of woody debris, capture of nutrients such as organic carbon in sediments during flood events, and a slowed movement of water through unconfined valley segments.¹ While abandoned meadows lack some of the complexity and characteristics of active meadows, similar process may still occur in meadows of both activity levels. This is because historic beaver activity leaves an impact on a meadow that gradually degrades over time.¹⁰ For example, Moraine Park, one of the abandoned meadows in this study, was noted in a 2012 study to be a site where beaver activity had greatly influenced post-glacial floodplain aggradation, despite having no beaver activity at the time of that study.¹⁰

The increased outflow abundance of the collector-filterer family Simuliidae and the collector-gatherer family Leptophlebiidae could suggest differences in availability of FPOM between meadow inflows and outflows. Simuliidae and other collector-filterers feed by capturing decomposing FPOM that is suspended in the water column as it travels downstream. Collector-gatherers such as Leptophlebiidae also feed on FPOM that has settled on the stream bottom.⁸ One possible mechanism for patterns of increased collector-filterers and collector-gatherers at the outflows is that river-floodplain meadows add FPOM to the water as it travels through meadows. This would lead to increased levels of FPOM at outflows, creating a more favorable habitat for the Simuliidae and Leptophlebiidae. Floodplain meadows are known to trap particulate organic matter during flood events when the stream is forced out of the channel and into unconfined meadow segments.¹ However, the sampling for this study was performed in August, after all sites had passed their peak flows and flood events. Another possible mechanism for increasing FPOM at outflows is the gradual breakdown of woody debris stored within floodplain meadows, continuously creating particulate organic matter.¹ It is possible that the slow breakdown of woody debris in these meadows could be providing an abundant food source for collector-gatherers and collector-filterers at the meadow outflows.

While collector-filterers and collector-

gatherers were more abundant at meadow outflows, the scraper mayfly family Heptageniidae, was found in greater numbers at meadow inflows. Since scrapers feed by grazing on periphyton attached to substrate, a difference in the amount of periphyton between inflows and outflows could be affecting the abundance of scrapers.⁸ Nutrients stored near the sediment water interface promote patches of periphyton, which are also influenced by stream geomorphology and oxygen availability.¹¹ Scrapers are also specialist feeders, which makes them more sensitive to disturbances or pollutants that impact a specific food source, unlike collector-filterers, which are more generalist feeders.⁹ Without further study, it is difficult to say which of many possible factors could be leading to lower Heptageniidae abundances at the meadow outflows.

Although technically statistically significant, results regarding differences between the inflow versus the outflow abundance of Dolichopodidae might not indicate anything about the impacts of land use and beavers. This family was only present in one location on one meadow, Mill Creek's inflow. It is possible that unique characteristics of Mill Creek's inflow are creating a favorable habitat for this family; these characteristics may or may not be related to the past land use and beaver activity in the downstream meadow.

Differences between active and abandoned meadows

Many aquatic insect taxa, including seven families and three entire orders, had higher abundance in active meadows compared to abandoned meadows, regardless of the location on the meadow. The families in **Figures 8 and 9** demonstrate examples of this type of relationship. Since inflows are assumed to be unaffected by the downstream meadow processes, one would expect to find differences between activity levels only when considering the outflows as they compare to inflows. However, since this is not the case, it appears that there are differences between active and abandoned meadows that are impacting stream communities at both inflow and outflow locations. There are several possible explanations for the overall increased abundances of certain taxa in active meadows.

For most of the aquatic insect taxa in this study, the adult forms are capable of flight. Adult aquatic insects are known to fly upstream from their larval habitat to lay eggs, resulting in upstream larval populations being supplemented by downstream populations.¹² If active meadow processes are creating particularly productive habitats at outflows, it is possible that some of the adults emerging from outflow habitats are traveling upstream to the inflow locations to lay their eggs.

Active meadow outflows may be a highly productive habitat due to the downstream flux of nutrients from processes such as the breakdown of woody debris and frequent flood events occurring within active beaver meadows.^{1,2}

A second possibility regards the idea that the presence or absence of beavers in these meadows is due in part to the meadow's land use history. Past land use could be influencing other overall characteristics of the streams and their riparian areas in addition to influencing whether the meadows are active or abandoned. An area's history could also be impacting the availability of food sources necessary for the presence of dietary specialists including scrapers such as Ameletidae, and shredders such as Nemouridae.⁹

While many families and orders were more abundant in active meadows, only two families were more abundant in abandoned meadows, Hydroptilidae and Hydropsychidae. Hydroptilidae are piercers that extract the contents of algae cells with specialized mouthparts.⁸ Algae may be more abundant in abandoned meadows than active meadows. When visiting these sites, it was observed that abandoned meadows sometimes had less riparian vegetation than active meadows. Less vegetation at abandoned meadows may be allowing more sunlight to penetrate the water and promote the growth of algae in abandoned meadows. Returning to sites and simultaneously sampling for algae and macroinvertebrates could help test this hypothesis. Meanwhile, Hydropsychidae are collector-filterers that construct a silk capture net which they use to strain out particulate matter. They are limited to particulate matter in a specific size range such that it won't pass through or break their capture net. It's possible that the streams with active meadows do not produce the correct size of particles for Hydropsychidae.

Interactions between meadow activity level and location

In this study, inflows of meadows of both activity levels were assumed to be equally unaffected by the processes occurring downstream in the meadow. When processes such as storage of woody debris and cycling of nutrients are unequal between active and abandoned meadows, there may be an interaction between meadow activity level and meadow location such that the relationship between inflow and outflow communities differs between active and abandoned meadows. The families in **Figures 10 and 11** demonstrate examples of this type of relationship.

One such interaction was observed in the scraper family Glossosomatidae, which was more abundant at outflows than inflows in active meadows, but not in abandoned

meadows. This suggests that processes occurring in active meadows are leading to more favorable outflow habitats for *Glossosomatidae* or for their food source, periphyton. One possibility is that increased periphyton thrive downstream of active meadows due to increased nutrients stored in the sediments of active meadows.¹¹ Nutrient rich sediments stored both in and just downstream of active meadows could be due to higher carbon accumulation and retention in beaver meadows through processes such as flooding.¹ It is also possible that competition for periphyton with other scraper taxa such as *Heptageniidae* affects the distribution and abundance of *Glossosomatidae*.

The collector-gatherer family *Leptophlebiidae* displayed the same interaction behavior as seen in the family *Glossosomatidae*. Collector-gatherers are one of the more generalist feeding groups of macroinvertebrates.⁹ They eat decaying particulate matter trapped in the sediments.⁸ Since active meadows are effective at trapping nutrients and particulate organic carbons in sediments during flood events, they may be contributing to downstream habitats with more nutrient-rich sediments, which would be a possible mechanism for the interaction behavior seen for *Leptophlebiidae* abundance.¹

Although not statistically significant, it is important to note that other taxa displayed a similar interaction behavior to *Glossosomatidae* and *Leptophlebiidae*. Two examples are the shredder family *Nemouridae* and the collector-gatherer family *Elmidae*. Similar to *Glossosomatidae* and *Leptophlebiidae*, both *Nemouridae* and *Elmidae* increased in abundance between inflow and outflow of active but not abandoned meadows. Shredders, such as *Nemouridae*, consume decomposing plant matter and wood, also called coarse particulate organic matter (CPOM).⁸ The gradual breakdown of stored woody debris in active beaver meadows could result in more CPOM at outflows. Also, shredders, like scrapers, are a feeding group of dietary specialists, which makes them more sensitive to disturbances and pollutants that could impact their food source than other feeding groups.⁹ Collector-gatherers like *Elmidae* are generalists that eat decomposing particulate organic matter trapped in substrates, which could be effected by meadow processes such as the trapping of nutrients. These insignificant interactions in *Nemouridae* and *Elmidae* could be due to a variety of differences between active and abandoned meadow processes or due to chance. Completing a similar study with more meadows and more replicates could be one way to investigate if there are interactions between meadow activity level and location in groups other than *Glossosomatidae*.

Dolichopodidae, *Uenoidae*, and *Perlidae* all displayed interactions as well. The interactions displayed by these three families involved a decrease in abundance between the inflow and outflow of the active meadow, and no change between inflow and outflow of the abandoned meadow. However, for *Dolichopodidae*, all six specimens found during this study were found at the inflow of Mill Creek, an active meadow. It is possible that characteristics unique to Mill Creek's inflow are creating a favorable habitat for this family; these characteristics may or may not be related to the past land use and beaver activity in the downstream meadow. Similarly, only three *Uenoidae* were found in this study and only in two places: the inflow of Mill Creek and the outflow of Hidden Valley. Again, since this family was only found in these two locations, it becomes increasingly questionable whether or not these patterns are related to river-floodplain meadow processes. Creating a statistical model that accounts for site-specific characteristics of the sampling locations could help determine if patterns in these rarely found families are significant to the processes of river-floodplain meadows. *Perlidae*, a predacious stonefly, displayed the same interaction as *Dolichopodidae* and *Uenoidae*, but was not as uncommon and was found repeatedly and at multiple locations. This predator may rely on some prey source that is more abundant at active inflows.

Conclusions and future actions

The use of aquatic macroinvertebrates as indicators of how anthropogenic disturbance affects stream quality has been repeatedly demonstrated by past studies.³ While the specific mechanisms impacting the abundances of different taxa in these meadows is uncertain, it is clear that these meadows do affect the aquatic insect communities. The results of this study suggest that both active and abandoned meadows are affecting the outflow communities, and that in some cases the processes in active meadows benefit certain outflow taxa more strongly. Also it appears that some characteristic of the active meadows or their surroundings is resulting in increased abundance of several taxa both upstream and downstream of the active meadows. The use of aquatic insects as a biological indicator in this study supports the findings of past studies, which have shown that important ecological processes happen in both active and abandoned beaver meadows.^{2,10} Both active and abandoned beaver meadows are ecologically significant areas and merit further study due to their importance to water quality and carbon cycling. Future studies could help link the macroinvertebrate community composition to the specific physiochemical and habitat characteristics of active and abandoned meadows. Since aquatic insects

are effective biological indicators, processes impacting them could be causing much greater impacts on the stream ecosystem and downstream water quality.

Acknowledgements

Thank you to Professor Tim Covino and Professor William Clements at Colorado State University, both whose advice and guidance were invaluable throughout this project. I thank Tristan Weiss for advice on project design and sampling in beaver meadows. Thank you to Tyler Lampard for assistance with fieldwork during benthic macroinvertebrate sampling.

References

- ¹Wohl, E., Dwire, K., Sutfin, N., Polvi, L. and Bazan, R. (2012) "Mechanisms of carbon storage in mountainous headwater rivers." *Nature Communications* 3.1. Pg 1-8.
- ²Wegener, P., Covino, T. and Wohl, E. (2017) "Beaver-mediated lateral hydrologic connectivity, fluvial carbon and nutrient flux, and aquatic ecosystem metabolism." *Water Resources Research* 53. Pg 4606-4623.
- ³Cairns, J. and Pratt, J. (1993) "A history of biological monitoring using benthic macroinvertebrates." *Freshwater Biomonitoring and Benthic Macroinvertebrates* 10.27. Pg 10-27.
- ⁴Poff, N., Olden, J., Vieira, N., Finn, D., Simmons, M. and Kondratieff, B. (2006) "Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships." *Journal of the North American Benthological Society* 25.4. Pg 730-755.
- ⁵Wallace, J. and Webster, J. (1996) "The role of macroinvertebrates in stream ecosystem function." *Annual Review of Entomology* 41.1. Pg 115-139.
- ⁶Merritt, R. and Cummins, K. (Eds.). (1996) *An introduction to the aquatic insects of North America*. Kendall Hunt.
- ⁷Downing, J. (1979) "Aggregation, transformation, and the design of benthos sampling programs." *Journal of the Fisheries Board of Canada* 36.12. Pg 1454-1463.
- ⁸Merritt, R., Cummins, K. and Berg, M. (2017) *Methods in Stream Ecology*, Volume 1 (3rd ed.). Academic Press. Pg 413-433.
- ⁹Cummins, K. and Klug, M. (1979) "Feeding ecology of stream invertebrates." *Annual review of ecology and systematics* 10.1. Pg 147-172.
- ¹⁰Polvi, L. and Wohl, E. (2012) "The beaver meadow complex revisited—the role of beavers in post-glacial floodplain development." *Earth Surface Processes and Landforms* 37.3. Pg 332-346.
- ¹¹Coleman, R. and Dahm, C. (1990) "Stream geomorphology: effects on periphyton standing crop and primary production." *Journal of the North American Benthological Society* 9.4. Pg 293-302.
- ¹²Müller, K. (1982) "The colonization cycle of freshwater insects." *Oecologia* 52.2. Pg 202-207.

Study Sites Located in RMNP

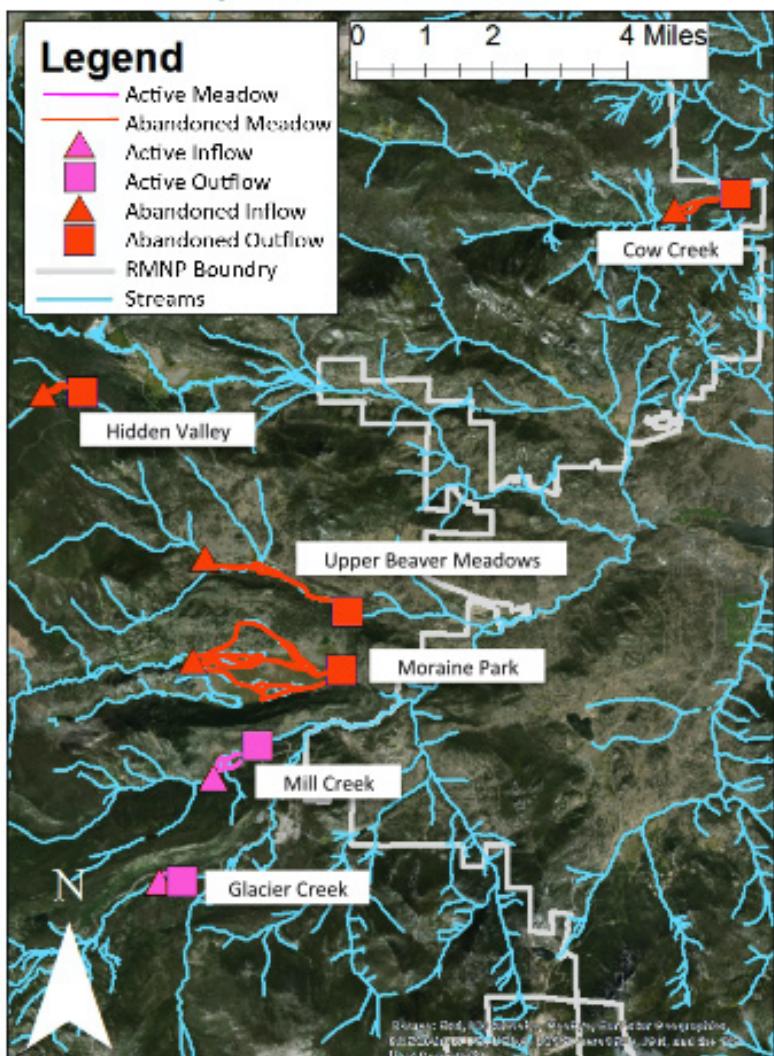


Figure 1: Study sites located in RMNP (left). Shows the 12 sampling locations at the inflows and outflows of the six river-floodplain meadows sampled for this study. All meadows are within the boundaries of Rocky Mountain National Park.

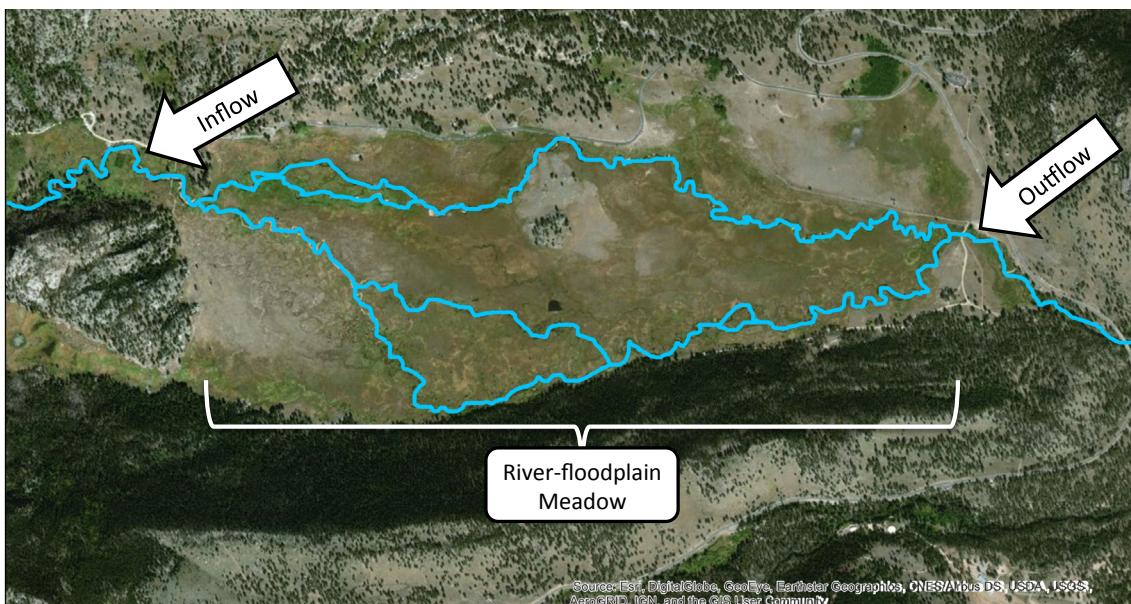


Figure 2: Meadow layout. A satellite view of Moraine Park, a river-floodplain meadow abandoned by beavers. The inflow and outflow sites are just above and below the section of stream that becomes braided as it travels across the meadow.

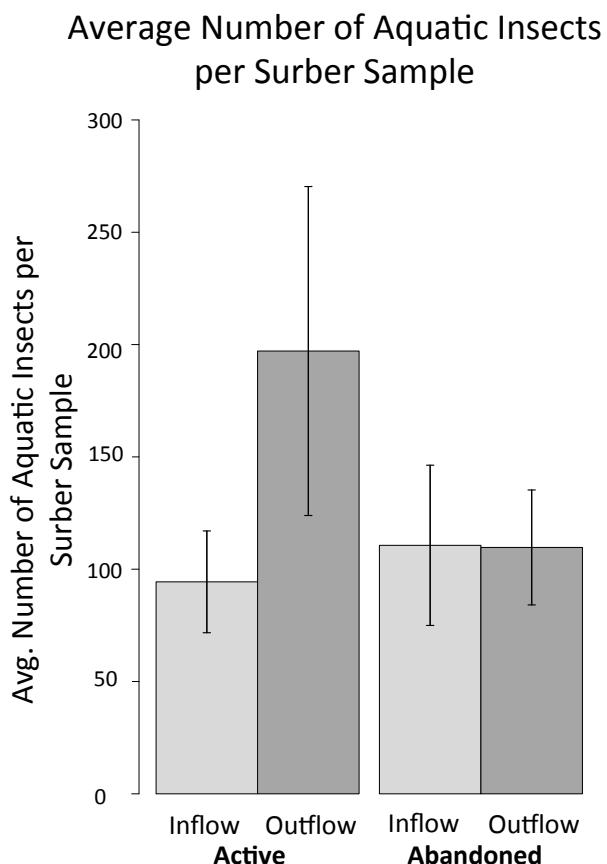


Figure 3: Average number of aquatic insects per Surber sample. This graph shows the average number of aquatic insects per Surber sample for each of the four treatment types: active inflow, active outflow, abandoned inflow, and abandoned outflow.

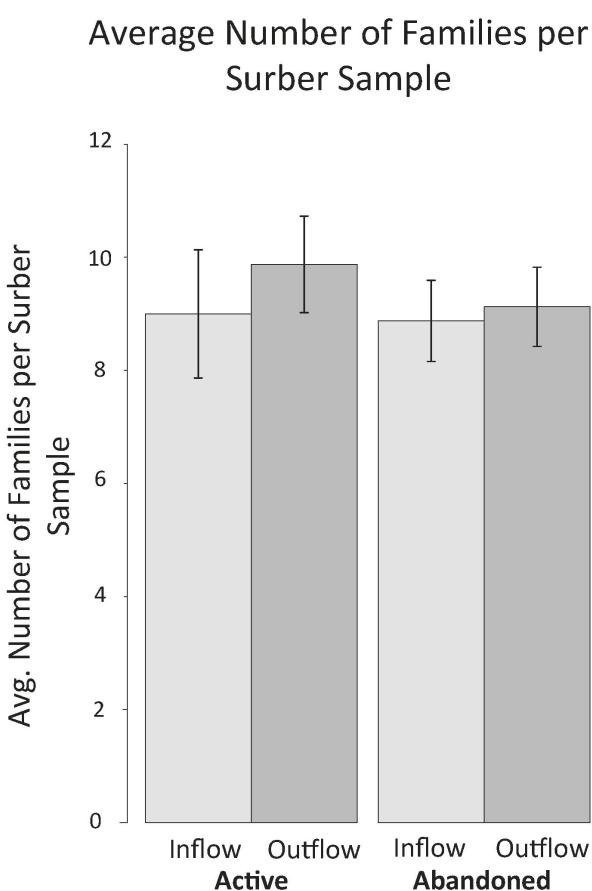


Figure 4: Average number of families per Surber sample. This graph shows the average number of different aquatic insect families per Surber sample found at each of the four treatment types.

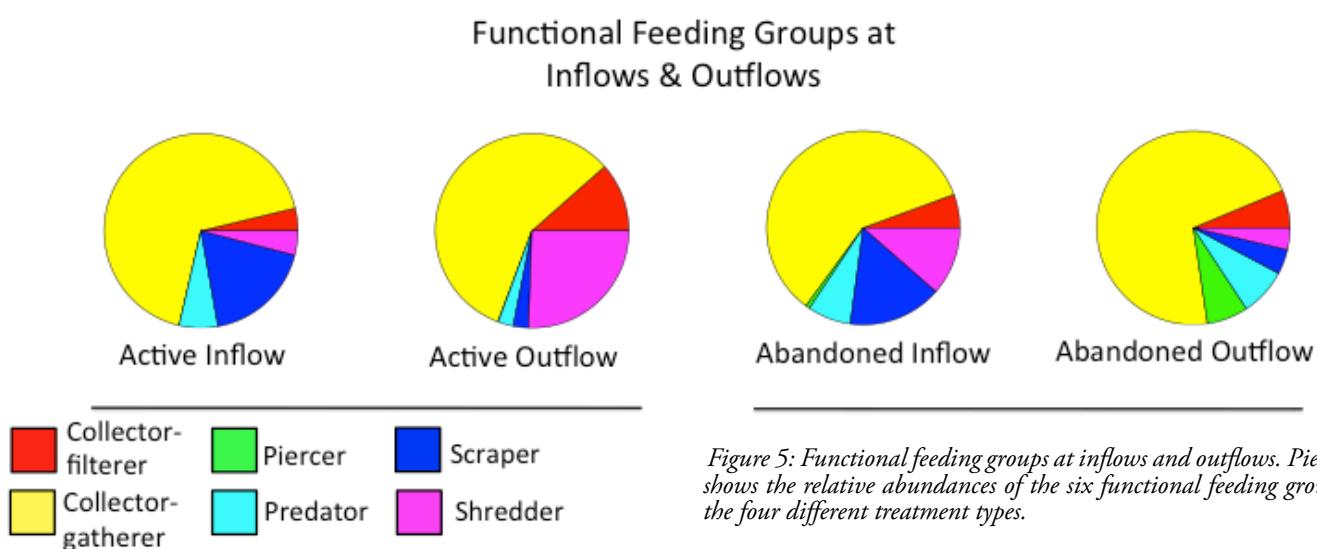
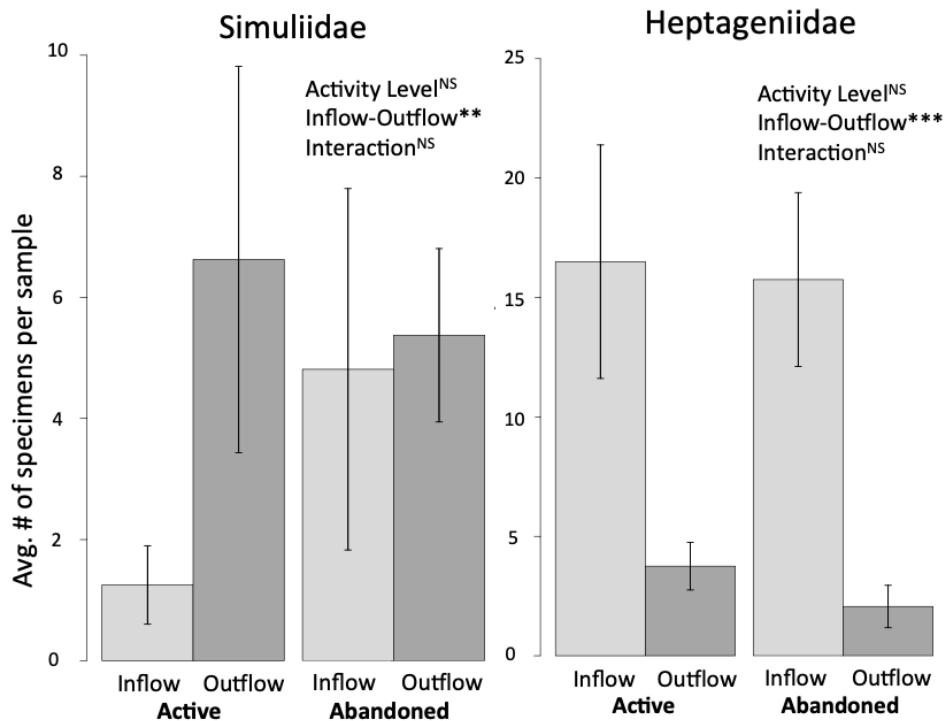
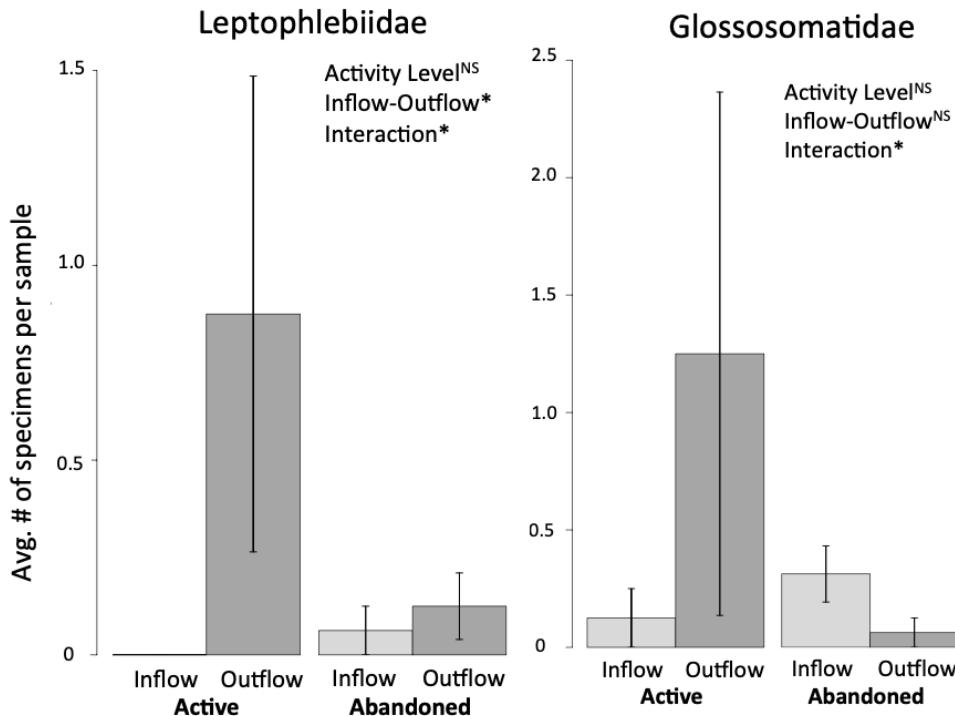


Figure 5: Functional feeding groups at inflows and outflows. Pie chart shows the relative abundances of the six functional feeding groups at the four different treatment types.



Figures 6 and 7: Examples of families that are significantly different between inflow and outflow locations. Figure 6 shows the average number of Simuliidae, a collector-filterer dipteran, which are more abundant at outflows than inflows. Figure 7 shows the average number of Heptageniidae, a scraper mayfly, which are more abundant at inflows than outflows.

Significance codes for p-values: NS > 0.1; 0.1 > * > 0.05; 0.05 > ** > 0.01; *** < 0.01



Figures 10 and 11: Examples of families that demonstrate an interaction effect between activity level and location. Figure 10 shows the average number of Leptophlebiidae, a collector-gatherer mayfly, at the four treatment types. Figure 11 shows the average number of Glossosomatidae, a scraper caddisfly, at the four treatment types. Both Leptophlebiidae and Glossosomatidae demonstrate an interaction effect, with their greatest abundance occurring at active meadow outflows.

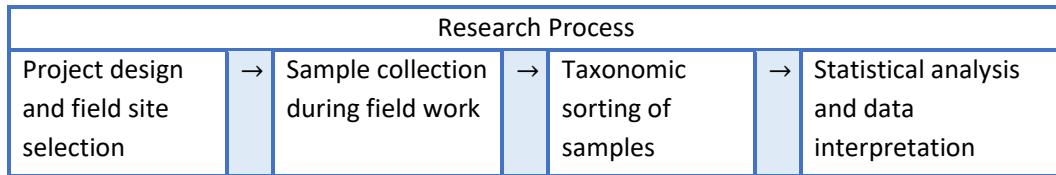


Table 1: Conceptual diagram of the research process. This diagram outlines the steps taken in completing this study.

Meadow Name		Sampling Locations:	Active or Abandoned?
Cow Creek		Inflow	Abandoned
		Outflow	
Hidden Valley		Inflow	Abandoned
		Outflow	
Upper Beaver Meadows		Inflow	Abandoned
		Outflow	
Moraine Park		Inflow	Abandoned
		Outflow	
Mill Creek		Inflow	Active
		Outflow	
Glacier Creek		Inflow	Active
		Outflow	

Table 2: Site Classification. There were twelve total sampling locations on six meadows. Each meadow was sampled twice, once at the inflow and once at the outflow. All meadows are classified as either Active beaver meadows or Abandoned beaver meadows.

Order	Family	Feeding Habit	Significantly differed in abundance across:		
			Activity Levels	Locations	Interaction: Activity x Location
Ephemeroptera					
	Ameletidae	Scraper	Active**		
	Baetidae	Collector-gatherer	Active**		
	Heptageniidae	Scraper		Inflow***	
	Leptophlebiidae	Collector-gatherer		Outflow *	Active outflow *
Plecoptera					
	Chloroperlidae	Predator	Active**		
	Nemouridae	Shredder	Active *		
	Perlidae	Predator			Active inflow**
Trichoptera					
	Brachycentridae	Collector-filterer	Active***		
	Glossosomatidae	Scraper			Active outflow*
	Hydropsychidae	Collector-filterer	Abandoned *		
	Hydroptilidae	Piercer	Abandoned *		
	Uenoidae	Scraper			Active inflow**
Coleoptera					
	Elmidae	Collector-gatherer	Active***		
Diptera					
	Dolichopodidae	Predator	Active**	Inflow*	Active inflow **
	Simuliidae	Collector-filterer		Outflow**	

Table 3: Results by family. The table shows the functional feeding groups of aquatic insect families from this study. The last three columns indicate if the family was found to be significantly more abundant in active or abandoned meadows, inflows or outflows, and if there was any interaction effect between the two factors.

1. Significance codes for p-values: NS > 0.1; 0.1 > * > 0.05; 0.05 > ** > 0.01; *** < 0.01

2. Families with only one total specimen found across all sites during this study were omitted from this list (Siphlonuridae, Dytiscidae, Dixidae, Ptychopteridae).

3. Families without significant ANOVA results were omitted from this list (Ephemerellidae, Perlodidae, Lepidostomatidae, Limnephilidae, Rhyacophilidae, Athericidae, Ceratopogonidae, Chironomidae, Empididae, and Tipulidae).

Order	Family	Abandoned Meadows								Active Meadows			
		Cow Creek		U.B. Meadows		Hidden Valley		Moraine Park		Mill Creek		Glacier Creek	
		Inflow	Outflow	Inflow	Outflow	Inflow	Outflow	Inflow	Outflow	Inflow	Outflow	Inflow	Outflow
Ephemeroptera		16	14.5	18.25	40	63.75	22.75	39.25	25.75	47	67.75	31.75	31
	Ameletidae	0.75	-	-	-	3	9.25	-	-	-	-	-	-
	Baetidae	5	7.75	3.25	36.5	35	5.25	6.25	12.75	11.5	45.75	22.5	23.5
	Ephemerellidae	6.5	6	5.5	2.75	4.25	7.75	4.5	6.25	11	17.25	0.75	2.75
	Heptageniidae	3.5	0.75	9.5	0.25	21.5	0.5	28.5	6.75	24.5	3	8.5	4.5
	Leptophlebiidae	0.25	-	-	0.5	-	-	-	-	-	1.75	-	-
	Siphlonuridae	-	-	-	-	-	-	-	-	-	-	-	0.25
Plecoptera		3.25	12	2.75	19.75	49.75	9.5	2.5	0.5	11.5	104.25	4.5	4.25
	Chloroperlidae	1.75	6	1.75	0.75	1.25	0.75	1	0.25	1.25	3	4.25	3.25
	Nemouridae	1.25	2.5	0.5	0.25	46	8.5	0.75	0.25	7.5	99.25	-	0.75
	Perlidae	-	2.25	-	-	-	-	-	-	1.5	-	0.25	0.25
	Perlodidae	0.25	1.25	0.5	18.75	2.5	0.25	0.75	-	1.25	2	-	-
Trichoptera		8.5	35	1	3.25	13.5	1.75	2	1.75	5.75	32.25	1.25	3.25
	Brachycentridae	1	2.25	-	-	-	-	0.5	1	3.5	31.25	1	0.25
	Glossosomatidae	0.25	0.25	-	-	-	-	1	-	-	-	0.25	2.5
	Hydropsychidae	3.5	2.5	-	-	-	-	-	-	-	-	-	0.25
	Hydroptilidae	3	29.5	-	0.25	-	-	-	0.25	-	-	-	0.25
	Lepidostomatidae	-	0.25	0.5	1.75	-	-	-	-	-	-	-	-
	Limnephilidae	-	-	-	0.75	0.5	0.75	-	-	-	-	-	-
	Rhyacophilidae	0.75	0.25	0.5	0.5	13	0.75	0.5	0.5	1.75	1	-	-
	Uenoidae	-	-	-	-	-	0.25	-	-	0.5	-	-	-
Coleoptera		6	6	9.5	1	1.25	2	4.5	10	18.5	57.75	4.25	5
	Elmidae	6	6	9.5	0.75	1.25	2	4.5	10	18.5	57.75	4.25	5
	Dytiscidae	-	-	-	0.25	-	-	-	-	-	-	-	-
Diptera		76.75	138.5	1.25	13.25	81	60.75	32.5	16	44.75	79.75	16.75	6.25
	Athericidae	0.25	-	-	-	-	1.25	-	-	-	0.5	-	-
	Ceratopogonidae	-	-	-	0.25	-	0.5	-	-	0.25	-	-	-
	Chironomidae	75.25	131.5	0.5	5	57	53.75	32.25	14	41	66.5	16.25	5.5
	Empididae	-	-	-	-	6.75	-	-	-	-	-	-	-
	Dixidae	-	-	-	-	0.25	-	-	-	-	-	-	-
	Dolichopodidae	-	-	-	-	-	-	-	-	1.5	-	-	-
	Ptychopteridae	-	-	-	-	-	-	-	-	-	0.25	-	-
	Simuliidae	1.25	6.5	0.75	7.75	17	5.25	0.25	2	2	12.5	0.5	0.75
	Tipulidae	-	0.5	-	0.25	-	-	-	-	-	-	-	-

Table 4: Averaged raw data. The table shows the raw averaged data, sorted by family. The abundance of each family is reported for each of the 12 sampling locations. Abundances reported here are the averages of 4 Surber samples taken per location. There were a total of 48 Surber samples taken during this study.

Canon and Blues: Diverse Influences on Philip Larkin's "Aubade"

BY BRANDON PAHL

UNIVERSITY OF IOWA

When considering writers who represent diversity, in terms of work as well as their life, one rarely encounters Philip Larkin. Especially after the publication of his collected correspondence, Larkin's reputation seems to have been cemented as that of a crotchety conservative, an upholder of traditional verse as well as traditional values. Such a way of viewing Larkin the man, it should be stated, is not necessarily incorrect (his letters abound, after all, in expressions of admiration for Margaret Thatcher, subtle jabs at modernist literature, and even some casual racism). When applied to his poetry, however, this way of viewing him is not only reductive and simplistic but downright incorrect. Much in Larkin's poetry suggests that he was not a complete literary conservative, but that, in crafting his unique and nuanced work, he drew on varied elements from canonical literature, modernism, and even the popular blues music of his time.

Although Larkin wrote four distinct volumes of poetry, and many uncollected poems, this analysis shall focus on one poem, "Aubade," which is arguably the most representative of his entire work. It came to fruition in 1977, near the end of Larkin's life, a period of nearly total creative stagnation. Thus, "Aubade" is perhaps his dreariest and most pensive poem, a bleak and candid reflection on the inescapability of death. As with many of his poems, it is at first glance a simple work with a simple theme; however, it reveals upon closer examination that he was not the typical, homogeneous English poet he is sometimes made out to be.

The Movement

Some (though not all) of the context needed to understand the origins of "Aubade" can be derived by understanding the literary movement Larkin belonged to, at least during his early career: "the Movement." "The Movement" is characterized as embodying "a general post-war period of reconstruction" in its approach to literature.¹ Movement poets sought a "return to traditional forms and regular structures of rhyme and rhythm," just as the English were seeking a restructuring of their society after the devastating events of the second World War.² This, supposedly, meant a turning away from the modernist aesthetics of pre-war American poets like T.S. Eliot and Ezra Pound in favor of newer, simpler verse (whether

or not Larkin himself completely avoided Eliot's influence in particular shall be examined later).

Larkin seemed to embody the Movement's aesthetic of simplicity not only in his poetry but in his humble working life as a librarian in the northern coastal town of Hull. He never had pretensions about being upper-class, and in fact seems to have been quite opposed to the stuffy, snobbish attitude that upper-class Englishmen (and especially Englishmen of letters) embody. His working-class sensibilities entailed a "lifelong enthusiasm for jazz," as well as an "irreverent rejection of the pieties associated with 'highbrow' arts."³ His personal artistic ethos was one of understandability: he wanted to write poetry that even England's uneducated laity could relate to and comprehend.

The very first line of "Aubade" is one that any working-class man, from any country for that matter, could perhaps relate to: "I work all day, and get half-drunk at night" (1).⁴ The rhythms and diction of the beginning of "Aubade" in particular are stark and simplistic, following strict iambic pentameter and consisting largely of monosyllabic words, thus maximizing its accessibility even to those who have no idea what the words "iambic" and "monosyllabic" mean. However, one should not be tricked by the gambit of simplicity at the poem's beginning, and thereby underestimate the ways in which "Aubade" is more than it seems. For, despite Larkin's publicly expressed wish that his poetry be simple and pure, almost as if it originated in a vacuum, much in this last great poem of his suggests the contrary.

Tradition and the Western Canon

The very title of "Aubade" goes against the philosophy that Larkin, and by extension the Movement, supposedly upheld. Rowe is quick to point out the French origin of the title, something that is rare for Larkin as he usually "took the official Movement line on foreign cultures;"⁵ namely, that to reference or use any language other than English was something of a *faux pas*. Sometimes the *mot juste* simply does not exist in English, which is ostensibly the case for "Aubade," a word that no one who has not studied literature extensively (whether they be upper, middle, or lower class) will fully comprehend.

Osborne defines an aubade in these terms: "a type of lyric in which (i) a narrator of one sex (ii) sings a love song (iii) to a member of

the opposite sex (iv) at dawn..."⁶ By this definition, Larkin's "Aubade" seems to be a traditional aubade only in the sense that it takes place at dawn.⁷ While it is not necessary to know the definition, etymology, and history of its title to enjoy or even understand "Aubade," possessing this information opens the poem up to further interpretation. The word "love" is used only twice in "Aubade," first in relation to remorse ("the love not given" (12)), second in explicit relation to death ("Nothing to love or link with" (29)), making Larkin's poem an ironic subversion of the traditional form.⁸ In Larkin's world, love is not only powerless in the face of death, but subordinate to it. Love is not even considered as a means of escaping or conquering death, whereas other means are considered and duly met by Larkin with acerbic skepticism.

Critical attention seems to be most often directed at "Aubade's" atheistic view of religion, a "vast, moth-eaten musical brocade," (23) but Larkin also examines and dismisses classic philosophical methods of confronting death, albeit in terms that the layman may easily digest.⁹ In addition to religion, "Aubade" also takes aim at two major Greek philosophical systems, Epicureanism and Stoicism, though neither is mentioned by name in the poem. Only the most basic ideas of these two philosophies are suggested and criticized, with the Epicurean viewpoint on death being the first to visit the examination table. "Aubade" finds "an aporia in the proposition that *'No rational being / Can fear a thing it will not feel...'*, a way of looking at death that is central to Epicurean philosophy.¹⁰ Larkin turns this sentiment on its head by arguing that "losing the means to feel is precisely 'what we fear'."¹¹ Speaking of death in these distanced terms is tantamount to an attempt to escape or deny it, making it little better than religion's explicit denial of death in favor of an afterlife. For Larkin, the sheer emotional horror of death should not be overlooked but faced—although we also cannot face it with a so-called "stoical" attitude.

With regards to Stoicism, "Aubade" is at odds with this philosophy in both its ancient and modern conceptions. Original Stoicism could indeed have passed as a religion or cult, and one of the main Stoic beliefs was "that death is not the losing of life but the returning of it to its Creator."¹² It may not fall into the category

of “moth-eaten musical brocade” (an epithet obviously aimed, most of all, at Christianity), but it still seems to be an idea that the speaker of “Aubade” would find suspect. Equally suspect is modern Stoicism, which is defined as an attitude of “courage or fortitude, indifference to pain or fear,” an attitude that hardly fits this poem’s speaker.¹³ The poem contradicts these sentiments with lines that are brimming with an almost sing-song mockery: “Being brave / Lets no one off the grave” (38-39).¹⁴ Courage in the face of death is something to be treated sneeringly. Any expressed lack of fear in the face of death is most likely a pretense; everyone possesses at least a modicum of said fear, and denying it is pointless. So much for philosophy. With “Aubade,” Larkin firmly maintains that the philosophies of antiquity, as revered and influential as they are, and with their notions of rationality and indifference, offer no true respite against the overwhelming terror of death.

As if incorporating and then subverting classical Greek philosophy in his poetry were not enough, Larkin does the same for the most revered English writer of all time: William Shakespeare. It needs hardly to be said that, at least when one writes in English, one cannot avoid Shakespeare’s influence—nor, for that matter, avoid comparison to him, especially when one’s verse becomes as well-known and potentially canonical as Larkin’s is. Larkin, naturally, was aware of all this. Many of his earlier poems use phrases and expressions culled directly from Shakespeare, though the material is often treated with a “schoolboy irreverence” as in the poem “Toads,” where “the desire to shout ‘Stuff your pension!’ to one’s employer is ‘the stuff / That dreams are made on.’”¹⁵ Larkin’s attitude towards Shakespeare could perhaps be seen as that of a sort of working-class jester who, while acknowledging Shakespeare’s indispensability, is still not afraid to poke fun.

Parallels may also be drawn between Larkin and the most famous of Shakespeare’s characters, Hamlet. These two are similar in that they can both be equally “melancholy and sardonic” and both have an obsession with death that can best be observed in Larkin’s “Aubade” and Hamlet’s “To be or not to be” soliloquy.¹⁶ However, Larkin still finds ways to differ from the greatest creation of the greatest English writer. “For Hamlet, Epicurus works; the thought of absolute annihilation cures his fear of death,” whereas for Larkin no such remedy exists, least of all in ancient Greek philosophy.¹⁷ In Larkin’s case, death is also explored with an intense, autobiographical candor (“Aubade”, whose first word is “I,” testifies to this), whereas Shakespeare creates a character who may or may not be autobiographical to think about death for him. Larkin’s poetic approach to death is unique; it is distinct, though perhaps not fundamentally different, from Shakespeare’s.

Modernism...?

Though Larkin’s own personal way of incorporating and subverting tradition in his poetry is unique, the very act of incorporation and subversion is not. To assert the contrary would be to ignore one of the most fruitful literary movements in history, modernism, which, in poetry, has arguably its greatest representative in T.S. Eliot. Larkin’s attitude towards Eliot seems to have been somewhat ambivalent, judging at least by his letters and critical writings, although he has had much to say about other notable representatives of modernism, one of whom is almost incessantly named alongside Eliot. He expresses a bald “dislike of Pound and Picasso,” in the introduction to *All What Jazz*, his collection of music reviews, and further inveighs against the entire modernist movement, as represented not only by Pound in poetry and Picasso in visual art but by Charlie Parker in music, by saying that “it helps us neither to enjoy nor endure.”¹⁸ (One wonders if the same argument could not also be made against “Aubade” itself.) In any case, Larkin’s expressed disdain for modernism would appear to preclude the influence of Eliot.

Yet dislike for a literary movement or artist is not enough to escape their impact, especially when they have been as impactful as Eliot has. The Movement may have designated Eliot the official status of *bête noire*, but unofficially his influence still seems to sift its way into the work of Larkin, supposedly the most representative poet of the Movement. Russell finds the most prevalent convergence of Eliot with Larkin in “Aubade,” arguing that the line “[t]he anaesthetic from which none come round... has clearly borrowed from Eliot’s vision of the evening ‘Like a patient etherized upon a table’ in ‘Prufrock.’”¹⁹ One wonders, however, if this image was so “clearly borrowed.” Has Larkin intentionally lifted this image from Eliot’s *second* most well-known poem, or has he simply come up with a similar (not the same) image unconsciously? Furthermore, if the former is the case, might this also be categorized as a “subversion,” in the vein of the more obvious, aforementioned subversions of canonical work?

In any case, an Eliot influence on Larkin, and especially on “Aubade,” cannot now be dismissed too easily. Their poetry shares not only extremely broad themes of death but narrower themes of the dehumanizing effects of modernity and city life—although the degree of subtlety with which Larkin and Eliot each approach these themes varies. Both “Aubade” and “The Love Song of J. Alfred Prufrock” concern themselves with these themes, but (and here Larkin’s distance from Eliot may be observed) they do so from opposite viewpoints. A younger man is clearly the speaker of Eliot’s “Prufrock,” a man who recognizes death but only as a remote certainty, embodied for him chiefly in the thinning of one’s hair. Although

the rough age of the speaker in “Aubade” is not so easily estimated, it seems hard to imagine death is as remote for him, especially in the lines “Not to be here, / Not to be anywhere, / And soon...” (18-20).²⁰ It bears repeating that Larkin wrote these lines when (though he did not know it at the time) he had less than a decade to live, whereas Eliot composed “Prufrock” when he was still young, with *The Waste Land* still ahead of him—a work that would appear to bear even less of an influence on Larkin.

Nevertheless, it is difficult to imagine a Larkin without Eliot, or for that matter, Larkin without modernism in general. That which he viewed negatively still had an effect on him, and in some ways his entire oeuvre including “Aubade” can be viewed as an answer to modernism (whether or not it is a *refutation* of modernism is contentious). One would be remiss however to focus mainly on things that Larkin himself viewed either negatively or ambivalently (some of which may have landed in his poetry only subconsciously), because, contrary to appearances, there were some things that he enjoyed and which influenced him. Larkin, like any adequate poet before him, did not find inspiration solely in the antiquated or even in the recent past, but also in the present, in the form of the popular jazz and blues music of his time.

Jazz and Blues

An appreciation for jazz had possessed Larkin from a young age, although his love for it is sometimes seen as peculiar. In his biography of Larkin, Booth writes that “[for] Larkin jazz was a private passion, shared with a small number of male friends. He never tired, in particular, of the ‘inexhaustible vitality of the blues’...”²¹ One would hardly expect Larkin, whose poetry in addition to his personal demeanor seems so saturnine, to be associated with something that possesses an “inexhaustible vitality.” And yet this phrase belongs to Larkin himself; it is culled from *All What Jazz*. Larkin’s admiration for jazz and blues was deep and sincere, and the influence of this music upon him may be observed not only in *All What Jazz* but in his poetic work.

The simplicity of certain types of jazz and blues songs especially affected Larkin’s poetry: “The twelve-bar blues formula, that modern version of the ancient aubade...gave him the example of a strict but infinitely variable artistic discipline: ‘for all its formal simplicity it is rarely monotonous’”²² “Formal simplicity” would appear to be the key phrase here, another quotation from *All What Jazz*, the introduction to which has already been acknowledged above as a key text for Larkin’s aesthetic philosophy, in poetry as well as jazz. His stance against jazz with modernistic tendencies, as exemplified by Parker, Davis and Coltrane, was firm. Jazz

after Coltrane “started to be ugly on purpose”; it was characterized by “chaos, hatred and absurdity,”²³ and while one could make the case that “hatred” can sometimes be found in Larkin’s poetry, “chaos” and “absurdity” seem mostly to be absent. Chaos and absurdity preclude accessibility for Larkin, which is arguably his central aesthetic tenet. The jazz and blues influences in Larkin’s poetry, rather than muddling the message of the poem, serve to make the message even clearer and more digestible for the average reader.

Leggett provides a detailed analysis of the blues influences in Larkin’s last great poem, arguing that “‘Aubade’ is not a blues, but...we can trace its strength to the blues.”²⁴ “Aubade” would, in other words, be a radically different (and probably less interesting) poem without the influence of blues music flowing through it. Leggett pays special attention to the first line of “Aubade,” with its drab iambic pentameter and its references to “[w]orking and drinking,” both of which “are...common properties of the blues.”²⁵ These coarse, simplistic terms are blues terms. A more traditional poem (a more *poetic* poem, one might even say) would certainly have found more eloquent ways of phrasing these subjects, if it decided to portray them at all. Larkin, however, is not concerned with being flowery, or with veiling the essential meaning of his lines. He, like many blues musicians, wishes to state the situation as matter-of-factly as possible. It is here that his poetry differs drastically not only from his modernist predecessors but from even earlier relatives like Hardy or Yeats, who most of the time refrained from such blunt expression.

Though the blues undeniably flows through “Aubade,” it is important to recognize that a white Englishman’s poem can only come so close to the true form. For as much as Larkin is concerned with being genuine, “Aubade,” if it can be considered blues, would be an inauthentic example of the genre, not only due to Larkin’s background, but also because it uses “an oral tradition within a written form.”²⁶ (Thus, even the Beatles’ attempts at blues are more authentic than Larkin’s.) The blues are meant to be sung aloud with plaintive, furious cries, not imparted to a sterile page. That the genre is a product of the long and painful struggle of black Americans for freedom and creative expression means that it can never be faithfully copied by someone such as Larkin. Nevertheless, Leggett makes the bold claim “that ‘Aubade’ is much closer to a blues tradition than to the long English tradition that preceded it.”²⁷ Larkin does not copy the blues, but, as with the works of English tradition, seeks to reinterpret them not just in terms of form but also of content.

Analyzing the content of “Aubade,” one returns to a consideration of its main and arguably only subject, death, which, it goes

without saying, is a wholly unoriginal theme in the grander context of world literature. Innumerable libraries could be stuffed with books of mediocre death-poems. “What is untraditional in [‘Aubade’],” what differentiates this poem from reams of dross “is its method of dealing with an anxiety that is both unacceptable and inescapable.”²⁸ This “method” involves complete sincerity, something which is undeniably derived from the blues, a form of music that is built around plain-spokenness. The lexicon of “Aubade” is largely simple, and when metaphors appear (which is rare), they are characterized by a distinct lack of whimsy. One might even criticize these few metaphors for being unimaginative: “The sky is white as clay” (48)²⁹ by no means scintillates with originality, nor is it intended to do so. This and other analogies in “Aubade” serve as crushing reminders of the grounded nature of death: it is a certainty of life that is as plain as “clay,” a word that could fit perfectly in the earthy vernacular of the average blues singer.

If such an analogy is not brimming with uniqueness, it is, however, easily comprehended. In no way does it betray Larkin’s professed desire to be genuine and understandable, to craft his poetry so that it is accessible even to the simplest layman—in short, to tell the truth. In this he has more in common with the blues than with most English-language poets that came before him, and this desire for truthfulness is found even in his earliest work. Who else, after all, would write a poetry collection entitled *The Less Deceived*, other than someone who wanted people to be “less deceived” by poetry, to find in poetry not inaccessible fluff but ideas and language that were close to their everyday experience? “Aubade,” as a later Larkin poem, is exceptionally in-tune with everyday experience, with its talk of postmen, wardrobes, curtains, et cetera. It possesses as well a subject that everyone can understand, as everyone must one day face death. “Aubade” “is as much about being *undeceived* as it is about death.”³⁰ The idea of being *undeceived* is one that Larkin *did not* derive from the English canon. Rather, he found it in the blues.

Conclusion

It is something of a commonplace that works of fiction and poetry that serve a blatant political or moral message tend not to last. Less of a commonplace (perhaps because it seems even more obvious) is the idea that work which either bears its influence obviously or, conversely, bears only one obvious influence, also tends not to last. The poetry of Philip Larkin generally, and “Aubade” as the possible crown jewel of his entire oeuvre, do not fall into any of these amateurish traps. Unless one looks at “Aubade” through a quite myopic lens, one inevitably finds not one, overwhelming influence but multiple diverse ones, which, although they make themselves apparent under

close examination, are, in the end product of the poem, kneaded together so seamlessly that the eye of the layman (as well, sometimes, as the professional) fails to notice them. Aubade, philosophy, modernism, blues: it is possible to see “Aubade” as both the synthesis of all of these things and the antithesis of them. All of these elements come together, under Larkin’s hand, to form something new: his “Aubade.”

References/Footnotes

¹Swarbrick, A. (2008) *Out of Reach: The Poetry of Philip Larkin*. Macmillan Pres. Pg 70.
²Ibid. Pg. 71.
³Ibid. Pg. 70.
⁴Larkin, P. (2012) *The Complete Poems of Philip Larkin*, edited by Archie Burnett. Farrar-Straus-Giroux. Pg. 115.
⁵Rowe, M.W. (2011) “‘Aubade’: Death and the Thought of Death.” *Philip Larkin: Art and Self: Five Studies*. Palgrave Macmillan. Pg. 167.
⁶Osborne, J. (2014) “Radical Citation: ‘Aubade’.” *Radical Larkin: Seven Types of Technical Mastery*. Palgrave Macmillan. Pg. 207.
⁷Even this could, however, be complicated by the poem’s first line, which implies repetition. “Aubade” may not describe one particular dawn but a great many which have come to morph together in the narrator’s mind.
⁸Larkin, P. (2012) *The Complete Poems of Philip Larkin*, edited by Archie Burnett. Farrar-Straus-Giroux. Pp. 115-116.
⁹Ibid.
¹⁰Osborne, J. Pg. 231.
¹¹Ibid.
¹²Ibid. Pg. 233.
¹³Ibid.
¹⁴Larkin, P. (2012) *The Complete Poems*. Pg. 116.
¹⁵Kerrigan, W. (1998) “Larkin and the Difficult Subject.” *Essays in Criticism: A Quarterly Journal of Literary Criticism*. 48.4. Pg. 294.
¹⁶Ibid.
¹⁷Ibid. Pg. 298.
¹⁸Larkin, P. (1985) *All What Jazz: A Record Diary 1961-1971*. Farrar-Straus-Giroux. Pg. 27.
¹⁹Russell, R.R. (2007) “Echoes of Eliot’s ‘The Love Song of J. Alfred Prufrock’ in Larkin’s ‘Aubade.’” *Heldref Publications*. Pg. 235.
²⁰Larkin, P. (2012) *The Complete Poems*. Pg. 115.
²¹Booth, J. (2014) *Philip Larkin: Life, Art and Love*. Bloomsbury. Pg. 317.
²²Ibid.
²³Larkin, P. (1985) *All What Jazz*. Pg. 21.
²⁴Leggett, B. (1999) *Larkin’s Blues: Jazz, Popular Music, and Poetry*. Louisiana State University Press. Pg. 95.
²⁵Ibid. Pg. 97.
²⁶Ibid. Pg. 99.
²⁷Ibid. Pg. 102.
²⁸Ibid.
²⁹Larkin, P. (2012) *The Complete Poems*. Pg. 116.
³⁰Leggett, B. Pg. 108.

J.R.R. Tolkien and Environmental Concerns in Mid-20th Century England

BY DREW HEIDERSCHEIDT

UNIVERSITY OF NORTHERN COLORADO

On December 5, 1952, a haze descended upon London. On average, visibility dropped to below 500m, but in the center of the city it was less than 50m. The streets were dark, empty, and enveloped in silence. People stayed in their homes, avoiding going out as much as they could. Five days later, the smog lifted, and 4,000 were dead, according to one British official at the time. The Great Smog of London, as it later became known, was the worst incident of air pollution in British history, and was caused by a variety of different factors, including air inversion and low-quality fuel, which was overall largely a result of a growing trend toward industrialization. Despite the scale of this disaster, air pollution and other forms of environmental degradation were nothing new to England, which alarmed many.¹ The alarm of those in England manifested itself in the artistic media of the time, notably in the works of J.R.R. Tolkien.

With the swift development of industry in 20th century England came a myriad of environmental problems such as deforestation and air pollution. As a result of the quick expansion of commerce at this time, people became more conscious of environmental degradation, something illustrated through the lens of literature and governmental regulation, such as The Clean Air Act of 1956. People generally considered the shift to mechanization in England detrimental, and many romanticized the times before any type of mechanization as a sort of “organic wholeness.”² Environmental concerns such as these are present in the works of the British philologist J.R.R. Tolkien, most famously the author of *The Hobbit* and *The Lord of The Rings*. In Tolkien’s works, there are a multitude of references to his environmental views, including the character of Saruman, a dark wizard who “sacrifices the value of permanence of sustainability for his grand scheme of domination and ownership.”³ Tolkien’s works were a representation of environmental worldviews in England during the 20th century, in which people were concerned with environmental degradation—particularly deforestation, air pollution, and industrialization. Not only this, but his characters reflected both industrialist and environmentalist mindsets.

Tolkien was fascinated by trees

throughout his entire life. He went as far as to call *The Lord of the Rings* “my internal tree.”⁴ Tolkien’s adoration of trees also led to an abiding fear of their mistreatment. “I am (obviously) much in love with plants and above all trees, and always have been; and I find human maltreatment of them as hard to bear as some find ill-treatment of animals,” he wrote to the Houghton Mifflin Company in 1955.^{5,6} It was the neglect of trees that led Tolkien to consider himself a guardian of forests. In a letter to the Editor of the *Daily Telegraph*, Tolkien wrote that he would defend all trees against their enemies.⁷ Tolkien’s fondness for trees is prevalent in his works, especially in the symbolic importance they take on. For example, the White Tree of Gondor in *The Lord of the Rings* is emblematic of the health of Gondor as a civilization. So, when the White Tree of Gondor is sick, Gondor as a whole is also in decline, and vice versa. Matthew Dickerson and Jonathan Evans observed Tolkien’s apparent fear of neglecting trees when they noted how Treebeard, a wizened animated tree in *The Lord of the Rings*, posited that cutting down trees for necessities is distressing enough, but unnecessarily tearing down trees is unacceptable.⁶ It is apparent that Tolkien considered the felling of trees immoral—so much so that the character of Treebeard acted as a mouthpiece for him to convey his thoughts on the subject. When Pippin, a Hobbit, asked Treebeard whose side he was on (i.e., good or evil), Treebeard replied he was on nobody’s side because “nobody cares for the woods as I care for them, not even the Elves nowadays.”⁸

Tolkien’s concern for trees in England was justified. The amount of forest felled from 1918 to 1945 totals up to 1,024,000 acres—over 3 percent of the total land area in England.² Tolkien mourned the loss of so many different trees in part two of *The Lord of the Rings*, *The Two Towers*. In response to the destruction of a large swath of forest, Tolkien wrote—through Treebeard—“Many of these trees were my friends, creatures I had known from nut and acorn; many had voices of their own that are lost for ever now. And there are wastes of stump and bramble where once there was singing groves.”⁹ The agony Tolkien felt when trees were destroyed was apparent, and his personification of

trees in *The Lord of the Rings* as individuals was indicative of their importance to him. Tolkien’s primary response to the obliteration of trees in the real world was the Old Forest. In it, Old Man Willow was the primary danger to Frodo—the story’s Hobbit protagonist—and his companions. The first time Willow appeared was on the first leg of Frodo’s journey to Mordor, the dark region ruled by the story’s antagonist, Sauron, when they traveled through the Old Forest. When the Hobbits (Frodo, Sam, Merry, and Pippin) took a stop to rest, Old Man Willow attacked them and nearly suffocated Merry and Pippin.⁹ However, the Old Forest and Old Man Willow were not inherently malevolent, as they appeared on the surface; rather, the “Old Forest was hostile to two legged creatures because of the memory of many injuries.”⁷ Thus, by virtue of trying to protect themselves from bipeds, the Old Forest and Old Man Willow were manifestations of the destruction of forests in the real world. Because in real life, “If someone starts hacking at a tree with an axe the rooted tree has to stand and take the blows,” this was Tolkien’s way of giving forests a defense mechanism in the face of environmental degradation.¹⁰

In addition to deforestation, air pollution was also a substantial problem in England during the 20th century. Factories in England released immense amounts of smoke into the air—something Tolkien observed in a letter to Rayner Unwin on October 24, 1952, when he wrote “This charming house has become . . . drenched with fumes,” and “I regret to note that the billowing cloud recently pictured did not mark the fall of Barad-dur [Sauron’s tower], but was produced by its allies.”¹¹ According to the environmental historian I.G. Simmons, air pollution in England at this time came in the form of “smoke from the quality of bituminous coal used, which was still about 40 million tonnes/yr.”¹² The sheer amount of smoke released into the air was recognized by James Law, a combustion engineer, when he said that in thirty years he had “never seen conditions worse than those in industrial cities and towns of the north during the past two years.”¹² There were major health concerns associated with the high quantity of coal being used in England at this time. In 1952, the English government

stopped rationing small-type coal, called “nutty slack”, and as a result, “Sulphur dioxide [reached] the highest concentration recorded since detailed monitoring of this compound had begun in 1932.” In the end, the Great Smog of London made possible by ubiquitous “nutty slack” led to the deaths of 4,000 people in England.¹³ In the aftermath, the British Parliament passed the Clean Air Act of 1956, which regulated the amount of pollution that could be put into the air.¹⁴ Tolkien himself felt the effect of air pollution on his own health, when he wrote in a letter to his son Christopher, “I have been afflicted with what may be . . . a ‘virus’ . . . of which the risk is steadily mounting in this polluted country.”¹⁵ To Tolkien, air pollution was ever-present, and impossible to ignore.

In one of the final chapters of *The Lord of the Rings*, Tolkien described air pollution in the native land of the Hobbits, the Shire:

There was a whole line of the ugly new houses all along Pool Side, where the Hobbiton Road ran close to the bank. An avenue of trees had stood there. They were all gone. And looking with dismay up to the road towards Bag End they saw a small chimney of brick in the distance. It was pouring out black smoke into the evening air.¹⁶

It is not by mere coincidence that Tolkien described the scene at Bag End in a way that highlights clouds of black smoke billowing through the air—it reflected his perspective that air pollution hindered England through its exacerbation of negative health impacts and interruption of the once bucolic aesthetic that the country held.

Industrialization, and Tolkien’s views on it, existed throughout his works. His views of industrialization’s effects are especially prevalent in the chapter “The Scouring of The Shire,” in *The Return of the King*, where Saruman forced industrialization upon The Shire. In a description of the Shire as it was forcibly industrialized, Tolkien wrote:

The great chimney rose up before [Frodo, Sam, Pippin, and Merry]; and as they drew near the old village across the Water, through rows and rows of new mean houses along each side of the road, they saw the new mill in all its frowning and dirty ugliness: a great brick building straddling the stream, which it fouled with a steaming and stinking outflow. All along the Bywater Road every tree had been felled.¹⁶

Tolkien’s description of the houses as “mean” and his characterization of the mill as “frowning” and “ugly” spoke to his opposition to industry and urbanization. It is unsurprising that Tolkien felt this way towards large industrial processes like iron or coal extraction, as the amount of land that was made unsuitable for any further

use was immense. By 1954, 173,000 acres worth of land in England and Wales had become unusable due to the iron and steel industries.² Indeed, in “The Scouring of the Shire,” Tolkien illustrated the problems of mills and industry. “[The mills are] always a-hammering and a-letting out a smoke and a stench,” Farmer Cotton (another Hobbit) complained, and added that they had polluted the nearby river.¹⁶ Likely, these mills were inspired by the ones Tolkien witnessed during his childhood.¹⁷ The extent of industrial development in England during the twentieth century was enormous, with sand and gravel production increasing from 127 million tons in 1950 to 192 million tons in 1960.² Tolkien bemoaned this very industry in “The Scouring of The Shire,” when he wrote “Great wagons were standing in disorder in a field of beaten bare grass. Bagshot Row was a yawning sand and gravel quarry.”¹⁶ Simply put, the Shire was almost entirely industrialized by Saruman. Even in Tolkien’s descriptions of Saruman, there was an inherent opposition to industry. Saruman “has a mind of metal and wheels; and he does not care for growing things, except as far as they serve him for the moment,” Treebeard complained in *The Two Towers*.⁸ For Tolkien, the threat of industrialization was forever looming on the horizon.

An interesting comparison may be made between the Shire at the beginning of *Lord of the Rings* to the one at the end. At the beginning, Tolkien described Hobbits as loving “peace and quiet and good tilled Earth: a well-ordered and well-farmed countryside was their favorite haunt. [Hobbits] do not and did not understand or like machines more complicated than a forge-bellows, a water-mill, or a hand-loom.”⁹ This was evocative of English sentiment before industrialization, when the most impressive technology was no more advanced than a hand-loom. By the end, the Shire was transformed at the hands of Saruman into an industrializing wasteland, blanketed by smoke and pockmarked with mills and quarries. However, the Shire was not defined by mills that pour out black smoke, but by vineyards, corn fields, woods, by vegetables like turnips and carrots, fruits like apples and plums, and flowers like snapdragons and daffodils.¹⁸

In *The Two Towers*, Tolkien painted a picture of the spread of industrialization through the use of the fortress of Isengard, home to the dark wizard (and one of the primary antagonists of *Lord of the Rings*), Saruman. In his speech about Sauron, Saruman, and Isengard at the Council of Elrond in *The Fellowship of the Ring*, the great wizard Gandalf evoked imagery reminiscent of

industrialization. “I looked on [the valley Isengard is located in] and saw that, whereas it had once been green and fair, it was now filled with pits and forges,” Gandalf said. He then added that around all of Saruman’s projects “a dark smoke hung and wrapped itself about the sides of Orthanc [Saruman’s tower].”⁹ It is obvious that Isengard was in the midst of a similar development to that of Britain in the 19th century. This industrialization is noted by Dickerson and Evans in *Ents, Elves and Eriador*—a book about Tolkien’s environmental worldviews—when they wrote that Isengard “is shown in the image of industrial smokestacks: vents and shafts spouting and belching fire and foul fumes.”⁶ Tolkien noted the harm of industrialization when he described the land around Isengard: “No trees grew there, but among the rank grasses could still be seen the burned and axe-hewn stumps of ancient groves.”⁸ From Tolkien’s description, the reader visualizes the stumps of trees, the putrid smells, and other traces of a place in the midst of industrialization. Treebeard described the situation, when he said, “Orcs came with axes and cut down my trees. I came and called [the trees] by their long names, but they did not quiver, they did not hear or answer: they lay dead.”⁸ This description of the pain caused by cutting down trees showed Tolkien’s anxieties of industrialism. Saruman personified an anthropocentric worldview—he used nature to increase his own power, at the expense of all living things.⁶ Tolkien juxtaposed the still-industrializing land of Isengard with that of the already-industrialized Mordor, Sauron’s abode.

Mordor, the home and fortress of the primary antagonist Sauron, is the most debased region in Tolkien’s legendarium, with the word “shadow” being regularly used as a descriptor. When Frodo and Sam arrived at the outskirts of Mordor, it was described as truly terrible. “[Frodo and Sam] had come to the desolation that lay before Mordor. . . A land defiled, diseased beyond all healing,” Tolkien wrote of the Dead Marshes.⁸ The fact that Tolkien used “desolate,” “defiled,” and “diseased” to highlight the status of the land that Frodo and Sam entered shows how badly damaged it was. Dickerson and Evans echoed this when they wrote, “The ground in Mordor is not only ‘fire-blasted,’ as with bombs, but also ‘poison-stained,’ as when factories’ toxic refuse contaminates the earth, water, and the air.”⁶ Tolkien’s description of the Dead Marshes (on the outskirts of Mordor) is telling:

[In the Dead Marshes] nothing lived, not even the leperous growths that feed on rottenness. The gasping pools were choked with ash and crawling muds, sickly white and

grey, as if the mountains had vomited the filth of their entrails upon the lands about. High mounds of crushed and powdered rock, great cones of earth fire-blasted and poison-stained, stood like an obscene graveyard in endless rows, slowly revealed in the reluctant light.⁸

This portrayal has a counterpart in Charles Dickens's descriptions of the industrialization of England during the 19th century, especially in his novel *Hard Times*. However, the word choice used by Tolkien here warrants further analysis. The fact that nothing lived in the area is reminiscent of the derelict land in Dickens's descriptions of industrializing England. And the land in the Dead Marshes (on the outskirts of Mordor) could easily be used to describe that same dilapidated land across vast swathes of English land. As a way to describe Mordor, Tolkien used imagery he had gained from "war, industrialization, and urbanism," to demonstrate how Sauron had destroyed the land.⁶ Tolkien expanded upon the desolation in Mordor when he illustrated the land surrounding Frodo and Sam in Mordor as "ruinous and dead."¹⁶

Part of the inspiration for Mordor was the First World War. Tolkien was a veteran of the First World War, specifically the Battle of the Somme, which was one of the largest battles of the war. By the end of the battle, both sides lost 1.2 million men combined, and by the end of the war as a whole over 10 million people died.^{16,19} "The Dead Marshes and the approaches to the Morannon," Tolkien wrote to L.W. Forster, "owe something to Northern France after the Battle of the Somme."²⁰ The Battle of the Somme, and other battles, considerably destroyed the forests of Europe. In the novel *The Somme*, A.D. Gristwood noted that "not a green thing survived the harrowing of the [artillery] shells. Constant barrages had churned the land into a vast desert of shell-craters, one intersecting another like the foul-pock markings of disease."²¹ Indeed, the First World War decimated the forests of France almost entirely, surely witnessed during his tour there. The landscape went from huge swathes of European beech trees, to almost total devastation.^{22,23} In the *Revue des Eaux et Forêts* during the First World War, one contributor remarked how not only the artillery was destroying the forests of France, but also the militaries there: "Vicious [military forest] exploitations' left stumps cut too low and felled plantations before they reached maturity."²⁴ Because of the war, the ecosystem of France (and other parts of Europe) changed entirely, with a diverse polyculture being turned into a monoculture dominated by Conifers.²²

Despite having never explicitly taken

the mantle of environmentalism, Tolkien imbued his best known creation—Hobbits—with strong environmentalist views. In the view of this author, environmentalism falls on a spectrum, ranging from passive to active. Those on the passive side support actions to protect the environment without taking action themselves, whereas those who are active take initiative to protect the environment. Ultimately, Hobbits fall on the extreme action side of that spectrum.

In a letter to Milton Waldman some time in 1950, Tolkien explained Hobbits collective environmentalism, when he wrote that Hobbits were "represented as being more in touch with 'nature' (the soil and other living things, plants and animals)."²⁵ The fact that Hobbits love the earth puts them in a similar category to environmentalists, who try to protect the earth. Both Dickerson and Evans took note of this when they wrote that "Hobbits are close to the earth, and they are closely associated with the material substance of the soil. They wear no shoes, and their walking around barefoot keeps them in direct physical contact with the earth," essentially meaning that Hobbits were closer to the earth than other creatures in Tolkien's legendarium because they were constantly in physical contact with it.⁶ From their introduction, it was clear that Hobbits were connected with nature and realized that they did not have dominion over the environment.²⁶

Tolkien was an avid gardener, and this was a trait that he projected onto Hobbits, particularly Bilbo Baggins (the protagonist of his earlier work, *The Hobbit*), Frodo, and Sam. In a letter to his son Christopher, he described his love for the garden. "When a gleam of sun (about 11) got through [the garden, it] was breathtakingly beautiful. . . I stood in the garden hatless and uncloaked without a shiver, though there must have been many degrees of frost," Tolkien wrote, in awe.²⁷ This close relationship with the environment was seen in environmentalists of the time, especially those who believed that rural England was one of "enduring values."² Throughout *The Lord of the Rings*, Hobbits have an inherent love for anything that grows. Bilbo and Frodo Baggins, both well-to-do Hobbits, never hired a housekeeper, a maid, or even a cook—and yet, they hired gardeners. And, if that were not enough, Tolkien named many Hobbits after flowers: think of the Hobbit women who appear sporadically throughout the series, all of them bearing names like Rose, Daisy, or even Marigold.⁶ Rather, Hobbits lived in harmony with their natural surroundings.

Of all the Hobbits, Sam Gamgee is the most prominent environmentalist—his profession is even gardening. In *The Return*

of the King, Sam believes Frodo is dead because he was stung by a giant poisonous spider, Shelob (although he was actually paralyzed). As a result, Sam took the One Ring. And, whereas Sauron or Saruman saw ultimate power when they had the ring, Sam saw only one thing: a garden so large that it could be its own country.¹⁶ In Tolkien's universe, this makes Sam the true hero of the story, and he wrote that it was "absolutely essential to the study of his (the chief hero's) character, and to the theme of the relation of ordinary life (breathing, eating, working, begetting) and quests, sacrifice, causes, and the 'longing for Elves,' and sheer beauty."²⁵ Indeed, Sam was representative of people's views in post-industrialized England: they missed the beauty and connection with nature of England's past.

When J.R.R. Tolkien began work on *The Lord of the Rings* in the 1937, after the publication of *The Hobbit*, environmental degradation was on his mind. Almost fifteen years later, upon publication, *The Lord of the Rings* represented environmental concerns in mid-twentieth century England. People were concerned about deforestation, particularly after two World Wars, which decimated England's forests; about air pollution, especially after the Great Smog of London in 1952; and industrialization, which many felt irrevocably changed the environment in England. Tolkien represented this in his descriptions of Middle-earth, filled with pristine environments, like the Shire, as well as degraded environments such as Mordor. Not only this, but Tolkien's characters often reflected both environmentalists (think: Hobbits, Treebeard, and others), and industrialists like Sauron and Saruman. In doing so, Tolkien created a microcosm that was reflective of environmental worldviews in 20th century England.

Acknowledgements

I would like to thank Dr. T.J. Tomlin for instilling in me the idea to combine Tolkien and Environmental History, and assisting me in the early stages of writing; Dr. Kristin Boivard-Abbo for her suggestions on different directions and sources, as well as for letting me borrow her extensive knowledge of J.R.R. Tolkien; and, Dr. Karen Barton for her unending support, editing, and numerous suggestions for this article over the past few years.

References

- ¹Brimblecombe, P. (2017) "London (1952): An Enduring Legacy." In *Air Pollution Episodes*. Brimblecombe, P. ed. UK Limited. Pg 57-72.
- ²Simmons, I.G. (2001) *An Environmental History of Great Britain: From 10,000 years ago to Present*. Edinburgh University Press.
- ³Elder, John. (2006) "Foreword." In *Ents, Elves*

LITERATURE

and Eriador: The Environmental Vision of J.R.R. Tolkien. Dickerson, M. and Evans, J. The University of Kentucky Press. Pg ix-xxi.

⁴Tolkien, J.R.R.(1962) Letter to Jane Neave, September 8-9, 1962. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 319-322.

⁵Tolkien, J.R.R. (1955) Letter to the Houghton Mifflin Company, June 5, 1955. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 217-223

⁶Dickerson, M. and Evans, J. (2006) Ents, Elves and Eriador: The Environmental Vision of J.R.R. Tolkien. The University of Kentucky Press.

⁷Tolkien, J.R.R. (1972) Letter to the Editor of the Daily Telegraph, June, 1972. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 419-420.

⁸Tolkien, J.R.R. (1954; repr. 1993) The Two Towers. Houghton Mifflin.

⁹Tolkien, J.R.R. (1954; repr. 1993) The Fellowship of the Ring. Houghton Mifflin.

¹⁰Flieger, V. (2000) "Taking the Part of Trees: Eco-Conflict in Middle-earth." In J.R.R. Tolkien and His Literary Resonances" Views of Middle Earth. Clark, G. and Timmons, D. eds. Greenwood Press. Pg 147-158.

¹¹Tolkien, J.R.R. (1952) Letter to Rayner Unwin, October 24, 1952. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 161-163.

¹²National Smoke Abatement Society. (1940) Smoke Abatement in Wartime. Copy Located at National Society for Clean Air and Environmental Protection, Brighton, UK.

¹³Thorsheim, P. (2006) Inventing Pollution: Coal, Smoke, and Culture in Britain since 1800. Ohio University Press.

¹⁴Clean Air Act, 1956, 4 & 5, chapter 52 (United Kingdom).

¹⁵Tolkien, J.R.R. (1971) Letter to Christopher Tolkien, June 2, 1971. In The Letters of J.R.R. Tolkien. Carter H. ed. Houghton Mifflin Harcourt. Pg 408-409.

¹⁶Tolkien, J.R.R. (1955; repr. 1993) The Return of the King. Houghton Mifflin.

¹⁷Carpenter, H. (1977; repr. 2000) J.R.R. Tolkien: A Biography. Houghton Mifflin.

¹⁸Sarjeant, W.A.S. (2001) "The Shire: Its Bounds, Food and Farming." Mallorn 39. Pg 33-37.

¹⁹Keegan, J. (1998; repr. 2000) The First World War. Vintage Books.

²⁰Tolkien, J.R.R. (1960) Letter to Professor L.W. Forster, December 31, 1960. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 303.

²¹Gristwood, A.D. (2002) "The Wasteland." In The Vintage Book of War Fiction. Faulks, S. and Hensgen, J. eds. Vintage Books. Pg 38-48.

²²Hupy, J.P. (2008) "The long-term effects of explosive munitions on the WWI battlefield surface of Verdun, France." Scottish Geographical Journal 122. Pg. 167-184.

²³Hupy, J.P. (2008) "The Environmental Footprint of War." Environment and History 14. Pg 405-421.

²⁴Pearson, C. (2012) Mobilizing Nature: The Environmental History of War and Militarization in Modern France. Manchester University Press.

²⁵Tolkien, J.R.R. (Undated) Letter to Milton Waldman. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 143-160.

²⁶Elgin, D. (1985) The Comedy of The Fantastic: Ecological Perspectives on the Fantasy Novel. Greenwood Press.

²⁷Tolkien, J.R.R. (1944) Letter to Christopher Tolkien, December, 1944. In The Letters of J.R.R. Tolkien. Carter, H. ed. Houghton Mifflin Harcourt. Pg 106-107.

“In the Public Interest”: The Historiography of Crusade Finance

BY LEE MORRISON

FLORIDA STATE UNIVERSITY

I. Introduction

Logistically, a medieval crusade presented many difficulties, particularly in regards to its financing. The act of raising an army, transporting it to the Holy Land, sustaining it upon arrival, and providing further support in the event that it succeeded in acquiring territory, was an enormous challenge that demanded considerable economic contribution. Crusading was incredibly expensive. Pope Innocent III and his successors were aware of this and realized that in order for any crusade to succeed, it had to be well funded. In 1208, Innocent declared a regional income tax on a tenth of the revenues of the clergy and laity in France to support the crusade against Cathar heretics in the south of the country. Innocent based this tax on the notion that “those acting in the public interest should be sustained by public taxes”.¹ Since 1199, Innocent had sought to find a solution to the inherent difficulties of organizing a successful crusade, especially by his push to introduce widespread clerical taxation as a source of funding. Scholars of crusade finance have traced the history of taxation and established a connection between the early tax initiatives of Henry II and the later adoption of the income tax as papal policy by Innocent III.

However, papal taxation was seen through a lens of cynicism by mid-twentieth century historians who considered the abuses of fiscal policy observed during the late thirteenth to sixteenth centuries as a natural development stemming from the programs of crusade finance established by Innocent III and his successors. Recent scholarship has suggested otherwise. While the relationship between secular and ecclesiastic taxation has proven consistent throughout the historiography, the nature of the papacy's crusade policy can be understood as being based in spirituality and practicality, as opposed to materialism. A central element of crusade policy was the establishment of taxation as a source of funding. Crusade taxation served a dual purpose. On one hand, it sought to provide a solution to the economic challenges created by crusading armies as a means of funding their transportation, supply, and manpower by providing financial support to the magnates who led their local forces on campaign. On

the other hand, it enabled the laity who were not fit for combat to participate in the crusade in a tangible way, contributing toward it on an individual level without actually traveling with the army. This was meant to ensure that those funds would go toward combatants only, eliminating the large portion of noncombatants who tended to follow the armies toward the Holy Land, as seen in the First Crusade.

Crusade taxation underwent considerable development in the thirteenth and fourteenth centuries, and though these taxes were often contentiously received, they did raise valuable funding. Though intended for crusade, privileges and taxes were eventually granted to secular rulers, who enthusiastically collected them but often without using the money which was raised for that express purpose. As secular rulers claimed the right of clerical taxation in their domains against similar assertions by the ecclesiastical government of the Church, the papacy in this period increased both the frequency of its taxation and the sources of its revenues. These abuses soured the popular attitude toward crusade taxation, not because of a loss of belief in the cause of crusading, but due to the constant levying of taxes, ostensibly for that purpose, which were then not actually carried out once the money was raised. The financial abuses which contributed to the Protestant Reformation, notably the sale of indulgences by the papacy, should be understood as having developed not out of the crusading policy of the thirteenth century papacy, but through later abuses by secular rulers and the Avignon popes.^A

II. Secular Precedents to Ecclesiastical Taxation

Papal taxation of the clergy was not an idea without precedent. Instead, it was inspired by secular initiatives of the twelfth century. Although Innocent III is credited as the first Pope to impose direct taxation upon the clergy in 1200, he was preceded by almost forty years of secular taxation which would continue to influence papal taxation after his pontificate. The levies of 1166, 1183, 1185, and most famously 1187 formed the basis of clerical taxation. This group of taxes will be referred to as the “twelfth century

secular taxes” for the duration of this section. These secular taxes proved to be models for later papal programs, most notably in their methods of collection and assessment. The levies set a precedent for the cooperation between ecclesiastic and secular authorities in crusade taxation, and its history is one of adaptation and flexibility. Innocent III's decision to adapt similar measures in 1200 was not borne out of a desire to squeeze cash from his clergy but was instead a product of practicality. He saw a longstanding problem (inconsistent and decentralized financing of crusades), and he saw a framework in place that already had mixed elements of lay and ecclesiastical participation and which had already produced successful collections, albeit at a local level in England. The decision to apply this framework to the Latin clergy and base it upon papal authority was a natural one.

The opening of the Vatican archives in 1881 made materials that would prove invaluable to the study of the medieval papacy available to students and scholars for the first time. Prior to their opening, no comprehensive treatment of the papacy's finances during this period existed. In 1934, William Lunt published what would become the first synthesis of medieval papal finance in which he examined the various sources of revenue available to the medieval papacy, though he made no mention of expenditures or repercussions of papal finance.² Expenditure would be treated in depth by a myriad of later historians. Lunt noted that papal taxation was influenced by twelfth century secular taxes and adopted their framework while simultaneously basing justifying extraction as an extension of the papal right of *plenitude potestatis*, by which the papacy could “assert its power over the Patrimony of the Church for whom the crusades were fought.”³ As papal taxes developed, so too, he argued, did the Curia, and the “mere organ of the [papal] household had become a great department of state.”⁴

Fred Cazel and James Brundage furthered the notion that the twelfth century secular taxes inspired crusade taxation during the pontificate of Innocent III, who wanted to expand and regularize them to apply to the papal patrimony at large.⁵ The twelfth century levies were the result of an agreement

between Henry II of England and the kings of France to tax the general populations of their realms to support the crusade.⁶ The use of Templars and Hospitallers, first appointed in 1185 to collect these taxes, lent them an element of ecclesiastic representation.⁷ Innocent wished to expand this ecclesiastic role while still cooperating with the kings to ensure that taxation went smoothly and without opposition. Cazel pointed out that the 1185 tax, which Lunt had previously considered to be nonexistent and based on a forged ordinance, was not only carried out but was instrumental in the development of the “Saladin Tithe” of 1188. Cazel argued that the 1185 tax was an example of the blending of secular and ecclesiastic authority present in the twelfth century levies. The tax was declared by Henry II and Philip Augustus with the consent of their bishops and counts, consisted of a three year tax of a hundredth on income and moveables (personal possessions which are not attached to the land), was backed by an indulgence granted by the bishops of Normandy, and was collected by members of military orders.⁸ Furthermore, the exemptions of necessities of profession from assessment and the terminology of the 1185 ordinance bore striking similarity to the ordinance of the Saladin Tithe.⁹ Brundage remarked that this sort of adaptation of previous taxes was characteristic of the financial system of the crusades.¹⁰

Ecclesiastical sanctions were used to enforce the twelfth century collections, further blending secular and church authority into the process of financing endeavors in the Holy Land. Henry II's 1166 tax, which called for two pennies per man for the first year and one thereafter, was to be collected in chests placed in each parish church, overseen by priests and enforced by sanctions.¹¹ This was developed further with Baldwin IV's 1184 levy of one bezant per hundredth of moveables, two bezants per hundredth of church, monastic, and baronial revenues, and a hearth tax of one bezant (or whatever they could manage) on the poor.¹² The revenues were to be collected by bishops.¹³

Cazel previously demonstrated that the 1185 levy was at a rate of a hundredth per moveables, and he concluded that this rate was borrowed from the 1184 levy and extended to cover a period of three years.¹⁴ Finally, the Saladin Tithe of 1187-88 was assessed at a rate of a tenth on moveables, with the same exemptions as 1185, but with the new stipulation that taxpayers assess themselves but pay before a mixed lay and clerical committee to ensure that the assessment was honest.^{15B} Not only were elements of ecclesiastic authority present in each of these taxes, but the Saladin Tithe

also saw the first attempt at ecclesiastic participation when Clement III requested that bishops contribute to the collections in addition to the laity.¹⁶ Cazel concluded that the Saladin Tithe offered the most promising basis for crusade taxation, but its successors were instead appropriated by secular princes and popes for purposes other than aiding the Crusader States.¹⁷

Though Clement III required that bishops should participate in the 1187 tax, Innocent III's declaration of a fortieth in 1200 represented the first general tax on clergy. In addition to the fortieth on clerical incomes, Innocent stipulated that he and the cardinals would pay a tenth of their own revenues.¹⁸ Clerical opposition to taxation was immediate and largely unceasing thereafter. This created collection problems. Whereas the English crown enjoyed a centralized authority within its territory and some prior history of implementing broad crusade taxation, the papacy possessed neither.

Innocent III attempted to use papal collectors in 1200, but clerical uproar made it clear that such a measure would be deeply unpopular. Cazel and Kay showed that Innocent and his successors were aware of this and did not initially favor outright centralization. Kay examined Innocent's attempts to balance local and centralized authority in 1200 and 1215, using local clergy to collect taxes under the supervision of Roman officials.¹⁹ He also adopted Henry II's use of parish chests in 1213 as a means of collecting alms for the Holy Land.²⁰ Powell also argued that Innocent was cautious about outright centralization, and attempted to learn and adapt from the organizational and operational failures of the Fourth Crusade. During that venture, the crusade army found itself impoverished and in considerable debt to Venice, culminating in the sack of Constantinople by Christian forces. According to Powell, Innocent was mindful of both clerical opposition and the failure of his financial plans in 1199 and sought conciliar approval for the 1215 tax instead of declaring it solely on his own authority.²¹

Collections for the Fifth Crusade were dreadfully slow and plagued by obstinate clergy, and by 1221, six years after the Fourth Lateran council declared a three year twentieth, it was still being collected.²² Kay argued that Honorius III, whom Lunt blamed for the turn to papal centralization, actually tried to avoid resorting to it and experimented with delegating legatine powers to four French bishops for collection of the Albigensian Tenth in 1221. However, he found that in doing so, the collection was hampered by personal vendettas and local feuds among those bishops, which

further delayed its completion.²³ The three Archbishops assigned with carry out collections each held their own provincial councils to decide on how to execute the collection within their jurisdiction, and they disagreed on the allowance of exceptions.²⁴ They convened in order to decide on the issue but deliberations were delayed by a jurisdictional dispute between Archbishop of Bourges and the Benedictine abbot of Deols, whom the Archbishop excommunicated according to his role as papal legate. The Archbishops proved overzealous in their execution of the collection, excommunicating so many clergymen for fraud that Honorius had to also grant them temporary power of absolution, as it was deemed excessive for the sheer number of excommunicates to journey to Rome for appeal.^{25C} Honorius turned back to using curial collectors not because he desired it, but because it was the only feasible method of collection, despite the clerical complaints.^D

France, Collins, Bird, and Asbridge furthered the work of Cazel and Kay. France, writing in 2005, found a precedent for both secular and ecclesiastic taxation in the intellectual and social advances of the twelfth century. He argued that the increase in wealth at this time caused society to turn to two institutions of authority, the papacy and the monarchies, to provide frameworks for law and stability.²⁵ After 1188, both institutions began developing greater administrative capabilities, leading to joint increases in the regularity of church and royal government, growth in education, and a general rise in intellectual development. Collins also pointed out an increasing focus on financial bureaucracy among the papacy and a trend of financially oriented popes. Cencio Savelli, before becoming Honorius III, compiled the first version of the *Liber censuum*, a record of papal assets and expected revenues.²⁷ Innocent IV was the son of an imperial tax collector, and he would raise the rates of taxation in the papal states.²⁸ Asbridge understood the Saladin Tithe in a manner which agrees with Cazel's claim that it represented the best model for funding crusades, arguing that it produced “the peak of a well-financed crusade army”, though he pointed out that only England had the administrative capabilities to enact such a tax.²⁹ Bird acknowledged the effect of the crusades as a stimulant on the overall economic atmosphere of the Middle Ages.³⁰

Overall, scholarship indicates that the twelfth century secular levies formed the basis for clerical taxation. This history can be traced to the levy of 1166 and the subsequent levies that were built into the Saladin Tithe, which became the model crusade tax. The first general tax on clergy came in 1200, but

elements of church authority were present in the previous secular taxes, and this blending of ecclesiastical and secular elements persisted in the policies of Innocent III. Issues arose immediately among the clergy and increased along with papal centralization of collections, though only after other experiments with local collection had failed. Monarchs and the papacy alike sought to cultivate their administrative capabilities, and taxation was a major venue for developing new articulations of authority and structures of governance. These experiments in fiscal organization developed as a response to the financial demands of crusades and the cost of maintaining subsequent possessions in the Holy Land.

III. The End of Secular and Ecclesiastical Cooperation

In the same manner that papal taxation grew out of secular innovations, the development of the secular state ultimately led to the devaluation and replacement of the clerical income tax. . Beginning in the late thirteenth century, competition between ecclesiastical and secular authorities over the right to tax local clergy led to popes levying more and more clerical taxes, decreasing their effectiveness while also cultivating more discontent. Lunt argued that secular rulers adopted clerical taxation to augment their own revenues, and as it became more profitable for them, it became less so for the papacy.³¹ Thus deprived of its primary method of funding crusade enterprises while simultaneously faced with more powerful and demanding secular administrations, the papacy turned elsewhere for financing any crusading endeavors.

Though he was focused more on the development of crusader legal institutions, Brundage did examine the use of crusade tithes by secular rulers. He noted that Edward I began a trend where English kings levied crusade tithes, but applied their own purposes and interests to the use of the revenues.³² In other words, money raised in the name of crusading was not supporting the Holy Land. Ultimately, Brundage argued that legal developments transformed the crusades into an internal European institution, and the popes increasingly used the “guise of the crusade” to raise money for political wars instead of Holy Land operations.³³ As the legal institutions of the crusade were expanded in canon law, secular princes sought to utilize crusader privileges for their own purposes.

Over a decade later, Housely would refute Brundage’s notion that the papacy was using the crusade as an excuse to launch political wars, but he continued to expand on the notion that secular states were encroaching

upon crusade taxation by wielding it for their own interests.³⁵ He demonstrated that the crusade tenth was used as a primary means of funding for the crusade against Frederick II’s son Manfred from 1255-1266, as well as for the initial phase of the War of the Sicilian Vespers. However, the tenth weakened after 1283 and continued to do so for at least twenty years with the successive defeats suffered by papal forces, many of whom were crusaders who had sworn legitimate vows.³⁶ At this point, popes such as Martin IV and Nicholas IV levied more tenths in order to offset the decline in value, but this proved fruitless.³⁷ By the time of Boniface VIII, tenths were being granted to allies outright in exchange for help in the Italian crusades.³⁸ Housely argued that this was a critical change which caused secular rulers to regard the clerical tenth not as a papal subsidy, but as a national asset which they could levy at will.³⁹ The growing power of the crown came to subjugate the clerical state, most notably when Philip IV, seeking funding for his wars with England, established royal hegemony over the tenth in his lands.⁴⁰ The Avignon papacy continued to levy tenths rigorously, but they were less important as a source of revenue and only caused more clerical dissent.

Boniface VIII attempted to curb secular appropriation of the tenth with the bull *Clericos laicos* in 1296, requiring papal permission before a secular state could tax church revenue. However, Phillip Stump pointed out that Clement V revoked the bull in 1306, and secular rulers had been levying tenths constantly since. Stump argued that reform of clerical taxation was one of the key points of contention at the Council of Constance and argued that secular rulers, previously thought to be the champions of church fiscal reform, actually opposed it, as they were already able to utilize clerical taxation to suit their needs by the time of the Council.⁴² Cazel also argued that political jealousies destroyed the cooperation between the papal and secular authorities on the issue of taxation. He pointed to the irony that, by the late thirteenth century, secular rulers levied crusading taxes for their own uses, yet they expected the clergy and the church to fund them when they did actually go on crusade.⁴³ The Saladin Tithe actually became the model of taxation in secular states, and the clergy were left as the primary economic losers of the entire crusading era.⁴⁴

Recent scholarship has only made the breakdown of papal and secular fiscal cooperation clearer. French monarchs were not the only secular authorities to appropriate church revenues to fund their wars. Joseph F. O’Callaghan considered Innocent IV’s concession of a portion of the *tercias*, a third

of the annual parish tithe, to Fernando III as significant in the history of the funding of the Spanish wars of the Reconquest.⁴⁵ He argued that in doing so, Innocent IV opened the door for Spanish kings to further levy ecclesiastic incomes for their own purposes.⁴⁶ John France pointed to the *cruzada* tax, an ecclesiastic tax granting minor indulgences which survived into the twentieth century in Spain, as a case-in-point example of secular bureaucracies laying claim to crusade taxation.⁴⁷

Just as scholars agree that the secular taxes of the twelfth century inspired papal taxation, they agree that in late thirteenth century, the growth of royal administrative capabilities was the impetus for its decline. Secular rulers turned to crusade taxes as a means of expanding their own revenues. The domestic crusades of the thirteenth and fourteenth centuries show a pattern of royal subjugation of the clerical state and a decline in the returns of crusade tenths. The Saladin Tithe, in which Cazel believed lay the key to successful crusade taxation, instead became the model of the secular income tax. In England, France, and Spain alike, kings laid claim to the revenues of the church in their jurisdiction, and the clergy emerged by the end of the thirteenth century with significant economic losses and a strong desire for reform.

IV. Crusade Finance as a Policy of Reform

While the connection between secular and ecclesiastic taxation has been fleshed out by decades of academic study on crusade finance, a more recent development reinterprets the nature of papal fiscal policy itself as it relates to crusading. Lunt and Brundage wrote that the developments of the legal and financial institutions of crusading were rife with abuses, and the motivations of crusaders and curialists alike came to be seen as materialistic.⁴⁸ However, beginning with Jonathan Riley-Smith in 1986, scholarship has indicated that spirituality, not material gain, must be recognized as a key motivating factor among crusaders to journey to the Holy Land. This method of evaluating the spirituality of crusading as opposed to the material aspects has since been applied to the crusade finance, and the resulting studies indicate that the thirteenth century fiscal policies associated with crusading were reformative in nature. It was the desire to maximize the practicality of crusading for all Christendom, not materialism, which ought to be recognized as a key motivator for the Curia in regards to clerical taxation. Abuses came about in practice, but the theory behind this taxation must be recognized for its insight on the motivations of participants in crusades.

As mentioned, Riley-Smith challenged the *status quo* of mid-twentieth century scholarship which held that the crusades were ultimately motivated by a desire for materialistic gain. He broke new ground by utilizing charter evidence to illustrate the various motivations of the first crusaders. He argued that crusading was, by nature, an expensive endeavor, requiring up to five times the annual knightly income.⁴⁹ Most crusaders went east without hope of financial gain, and those who returned carried little wealth back with them.⁵⁰ What spoils were gained were spent almost immediately on food and equipment, which became ever more scarce and costly as the expedition pushed deeper into the Levant.⁵¹ He also disproved the notion that families used the crusade as a platform to offload their greedy second sons. Riley-Smith pointed out that families made considerable financial sacrifices to send their sons, husbands, and fathers on crusade, mortgaging their own inheritances and receiving little to no profit as a result.⁵² Recent scholarship has examined papal financial policy through this emphasis on spiritual motivations.

Riley-Smith's point on families making financial sacrifices from the outset of crusading indicates that when Innocent III issued *Quia Maior* in 1215, famously calling on every member of Christendom to take the crusade vow in order to either carry it out physically or commute it through donation, he did not do so without precedent. *Quia Maior* did not invent the notion that non-combatants should contribute financially to the crusade; it simply formalized an idea long in gestation among the Curia.⁵³ Non-combatants had already been contributing financially to the crusade since 1095 and were doing so without enjoying any crusade privileges or rewards in return. France argued that Innocent was not pursuing cynical fundraising, but was trying to extend the salutary effects of crusading to the laity as a whole who had already been involved in bearing the economic burdens of crusading since 1095.⁵⁴ He posited that Innocent saw financial support of the crusade as a purifying act, even if it was impersonal.⁵⁵ Commutation of vows was thus seen as a practical extension of crusading privileges to all Christians, a dual effort on Innocent's part to foster support for the crusade while also offering salvation to as many souls as possible. Innocent's successors continued to refine this policy, culminating in Gregory IX's bull *Rachel suum videns*. In his study of the Baron's Crusade (1231-1241), Michael Lower also considered the idea of Christian unity through crusading to be the primary policy of the thirteenth century popes.⁵⁶ Gregory IX, he argued, sought to build off

of Innocent's programs and muster a "mass mobilization of purses" through redemption of vows.⁵⁷ Though Lower's study proved that this policy ultimately failed for Gregory, the belief that a willing laity could be united in participation through a wider availability of the crusade indulgence can be seen as a theme of crusade taxation throughout the first half of the thirteenth century.⁵⁸

Bird, Bolton, and Parker have contributed greatly to the understanding of the reformative nature of Innocent III's policies. Innocent sought to recover lost revenues and generate new ones in kind through a reform of both curial and local parish finances. Crusade and fiscal policy were interconnected for Innocent, and he sought to use both to return his Church to good standing, utilizing vow redemptions as a method of extending salvation to as many souls as possible.^E Innocent also consolidated curial finances, demanding proper census payments and establishing fixed fees for curial services instead of allowing officials to simply collect gifts. Furthermore, Parker noted that Innocent focused on establishing effective collection while also legislating against abuses by collectors. *Pium et Sanctum* required crusade preachers to minimize their entourages, refuse personal gifts, and live frugally so as to embody Christian poverty.⁵⁹ Ostentatious preachers, the pope concluded, were not likely to collect alms from the faithful, as they did not embody the image of the Church which Innocent wished to project.

Innocent also sought to stabilize local parish finances and protect parochial tithes by eliminating tax shelters, ending exceptions on land acquired by religious orders, and assessing tithes before other dues to ensure that parishes saw increased revenues.⁶⁰ Bird and Parker argued that Innocent's motive in doing so was to provide more revenues for local communities, hopefully making them more likely to participate in taxation. They concluded that Innocent was mindful of public opinion regarding the church and sought to make it more positive, primarily through financial reform in order to stimulate crusade involvement.⁶¹ If parishes had more revenue in general, they would be more willing to contribute a portion of it to crusade by commuting their vow as an act of participation, especially if a partial indulgence was extended to them for doing so. Innocent's fiscal policy was thus dually aimed at facilitating the crusade while simultaneously offering salvation to his flock of souls.⁶² Curial and parish finances were reformed in order to stimulate a healthy public perception of the church, and the turn to vow commutation was the most practical way of involving the laity in crusade. It was

not, as thought by Lunt and Brundage, a purely materialistic act.

Though Riley-Smith initially focused on the relationship between spirituality and crusader motivations, it is logical to examine papal fiscal policies in this way as well. By examining charter evidence, Riley-Smith also illuminated the substantial sacrifices that noncombatants and families were already making to crusading, even from its outset. This formed the basis for Innocent III's initiative of mass signing and vow commutations. When this is applied to the image of the early thirteenth century papacy as presented by France and Lower, it could be suggested that such acts as *Quia Maior*, previously seen as cynical fundraising, were actually practical extensions meant to both foster support for the crusades and maximize their salvific effects. Scholars have recently interpreted Innocent's policies as being reformative in design and motivation, meant to improve curial finances and cultivate parish revenues while also addressing known complaints over revenue collections. Whereas earlier scholars such as Brundage and Lunt saw these developments solely as extensions of curial power, Bolton, Bird, Smith and Parker have shown that their basis was more spiritual and practical. Innocent's fiscal policy in relation to crusading was grounded in what he perceived as his responsibility to save as many souls as he could.

V. The Council of Constance and Indulgences as a Revenue Source

The idea that indulgences were intended as a source of revenue for the Church is noticeable throughout early twentieth century scholarship, likely derived from contemporary critics of ecclesiastical wealth such as Matthew of Paris. However, more recent study has indicated that this is not the case. Innocent III's decision to install mass signing and vow commutation was not cynical fundraising, as mentioned previously. The work of historians such as Stump and Bird has shown that indulgences were not sold to generate revenue themselves but meant instead to reward donations and enable widespread participation while also ensuring the military viability of the army.^F It was only after the devaluation of the crusade tenth, the widespread taxation of the Avignon papacy, and the reforms of the Council of Constance in 1418, that indulgences came to be treated as "spiritual revenue". Indulgences became revenue only after the program of papal taxation was handicapped.

As mentioned, Lunt treated indulgences as a source of revenue, and he interpreted the tendency of seekers of papal pardon to make offerings in exchange for commutations of

their vows as evidence of material motivation and papal corruption.⁶³ These offerings normally found their way into local church institutions or were sent off with crusaders, but by the 15th century, all receipts from indulgences went to the papal Camera.⁶⁴ What Lunt neglects to mention is how this transition occurred. Brundage made a similar argument, pointing out that the failure of the second crusade made it obvious that a method of participation for the militarily unfit was needed. However, Brundage believed that this decision was motivated by the desire to turn enthusiastic noncombatants into a source of income, but he did not consider the spiritual implications of doing so.⁶⁵ Bird, Bolton, and Parker illustrated that spirituality was very much present in the minds of Innocent III and his successors as they made vow commutation into regular policy. Lay non-combatants had been participating in the crusade movement from its beginning, both as active pilgrims and by making substantial economic sacrifices to send family members on crusade. They were involved in financially supporting the crusades from the beginning, and the extension of an indulgence to them in exchange for vow redemption can be seen as a means of practical policy that fell in line with Innocent's desire for widespread involvement and widespread salvation.

Stump saw the Council of Constance as the point from which indulgences became a source of "spiritual revenue". He challenged the notion, held by historians such as Hübler and Valois, that Constance did not lead to reform, and that curialists circumvented the desire of the French and German representatives.⁶⁶ The Council reformed both the process of levying tenths and other methods of papal income, notably spoils, annates, and services, which by that point had overtaken tenths as the primary source of papal revenue.^{67G} The German reform party wanted to require approval of a general council before a tenth could be levied, and Hübler believed this to be an indication that Constance represented a movement toward conciliar authority, resulting in escalating anti-ecclesiastical tensions when this was denied.⁶⁸ Stump indicated that only a small amount of reformers actually wanted conciliar authority, and even the German king Sigismund, thought by Hübler to be a reformer, would have opposed this move.⁶⁹ The general council would have met once every ten years, and requiring its authority to levy a tenth would have undoubtedly destroyed the effectiveness of the tax. Instead, Martin V settled on requiring consent of territorial clergy for local tenths and the consent of the College of Cardinals and prelates for universal tenths.⁷⁰ Stump argued that this was indeed a considerable

success for the reformers. It required not only that the Pope notify and meet with the clergy, but he also required their subscription before declaring a tenth. Further, a universal tenth had not been levied since 1343; instead localized territorial tenths were consistently raised by the Avignon papacy.⁷² These changes demonstrate the growth in representative bodies in response to taxation, but they did not indicate a wholesale rejection of the validity of crusade taxation. Though participation was still willing, representative bodies desired a role in dictating the terms of potential taxes.^H

Stump contended that these changes actually limited papal power and effectively ended crusade taxation as an important source of income. As previously evidenced by France and O'Callaghan, the *cruzada* tax continued, as did others, but these were insignificant. Crusade taxation had been declining in value for the papacy since the late thirteenth century, and the more significant sources of revenue (spoils, annates, and services) were constantly reserved by Avignon popes to the chagrin of local clergy. This practice was also reformed at Constance, as evidenced by Stump. Martin V ended papal reservations of spoils and procurations and left them to local clergy, and he extended the payment of services by an additional year while also reducing the assessment itself by half.⁷³ This was monumental. Services, especially in France, were a critical part of papal income, and the reform of Constance resulted in major losses to papal revenue.⁷⁴ It was only after these losses, Stump argued, that the papacy turned to "spiritual revenues" (indulgences) and pressed their seigniorial claims in Italy and became even more embroiled in Italian politics.⁷⁵ It was only after Constance that the outright sale of indulgences became papal policy.

According to Collins, the clergy were always opposed to taxation, and as the Avignon papacy continued to levy tenths at an increased pace, reserve annates and spoils, and clamp down on payments of services, this opposition only increased.⁷⁶ Thus, coinciding with Stump's study, it is clear that the Council of Constance was not meant to be a precursor to the Protestant revolution by attempting to instill a sort of parliamentary authority over the church. Instead, reformers were simply fed up with abuses from papal and secular collectors alike, and they sought only to limit the burden of taxation upon themselves. In doing so, they created a vacuum in papal sources of revenue which caused the outright sale of indulgences to reach its apogee, and indulgences took on a nature differing from their crusade financing applications.

Bolton, Bird, and Parker made it clear

that indulgences were never intended as outright revenue by Innocent III or his thirteenth century successors. The policy which produced the crusade indulgence was not the same as the policy which produced the indulgences which Luther denounced. The involvement of crusade indulgences in financing of crusades was meant to incentivize the laity to take the cross, give alms, and pay taxes. The use of indulgences as incentives was also not without precedent, as Eugenius III had offered a 1/7 indulgence to those who donated to the Second Crusade, as noted by Collins and Phillips.⁷⁷ There was no systemized taxation at this point, and payments by the laity toward the crusades were still entirely optional. Innocent was acting on precedent. Gregory VIII gave indulgences in return for donations after the fall of Jerusalem in 1187, and the strength of the indulgence was proportional to the amount given.⁷⁸

It is deceptively easy to see this as cynical materialism as Lunt and Brundage did, but recent studies have raised a powerful alternative interpretation. Bird and Parker analyzed Innocent III's education in Paris under Peter the Chanter before he became pope, which sheds valuable light on how Innocent handled indulgences. Peter the Chanter did not approve of indulgences in theory, but he recognized their practicality to sinners who have no other options.⁷⁹ Innocent considered them in this same vein and treated them as a minor spiritual cost which would result in a significantly greater spiritual benefit. The indulgence of the twelfth and thirteenth centuries was not thought of as revenue, but as the most effective way to involve the laity in the crusade. It was only after crusade taxation was claimed by secular powers, abused by the Avignon papacy, and reformed at Constance that the indulgence changed from a practical spiritual incentive to an outright source of revenue.

VI. Conclusion

By analyzing the scholarship of crusade finance, several trends can be identified, some continuous and some recent. Scholars assert that secular taxation in the eleventh century was built on adaptation, and subsequent taxes were influenced and shaped by previous ones while also incorporating elements of secular and ecclesiastic authority. The success of the Saladin Tithe inspired Innocent III to turn clerical taxation into papal policy in 1200, and here again subsequent taxes were built on the successes and failures of previous ones. The historiography is consistent on this point. The papal approach to crusade finance in the early thirteenth century was motivated by practicality and spirituality. A

growing scholarly interpretation now exists in which the expansion of papal financial policy by Innocent III and his successors through taxation and vow commutation was simply the most feasible way to steer support crusading efforts and extend salvific benefits to the laity. These taxes were thus seen by the papacy as a tool of pastoral care. Fiscal policy and crusade were therefore tied together in an atmosphere of reform meant to touch every corner of Christendom.

As secular and curial administrations matured, cooperation turned to competition. Kay proved that curial centralization of collections was not the outright goal of Innocent III or Honorius III, and it only became the norm when other experiments involving local authority failed. Thomas Smith has advocated for a reinterpretation of the maligned career of Honorius III which highlights the responsiveness of his diplomacy, and his attempts at crusade taxation were likewise an authentic attempt to respond to external challenges.⁸⁰ The appropriation of crusade taxation by secular authorities is also consistent in the historiography, but until Stump's study, not enough attention was paid to the effect that the Council of Constance had on papal finance. Stump reversed the assumption that the failure of the conciliar movement at Constance was a precursor to the Protestant Reformation. Instead, he indicated that papal finances were effectively handicapped by the Council's reforms, and it was only after losing nearly every other available source of revenue that the papacy turned to the outright sale of indulgences. Prior to this, indulgences were not seen as sources of revenue, as Bird and Parker pointed out, and Innocent and his immediate successors treated them as a practical way of extending the salvific effects of the crusade to the laity.

The abuses which were reformed by the Council of Constance were not explicitly connected to the fiscal programs developed in the early thirteenth century by the papacy. A line of connection should not be drawn between Innocent III's and his successors' use of indulgences for funding crusades and the liberal distribution of indulgences adopted by Leo X. Far from being a cynical cash grab, the financial programs instilled by the papacy in the early thirteenth century were built on a mix of spirituality, practicality, and reform. The intention was to involve as many Christians in the crusade as possible, not for the good of the papacy, but in the public interest of Christendom as a whole. Just as Riley-Smith's focus on spirituality fostered a new movement in crusader studies, a continued examination of the spiritual motivations and justifications behind the methods of crusade finance employed by

the papacy can foster a new understanding of what these policies were truly intended to achieve.

VII. Footnotes

^AWhether the fiscal policies of the Avignon Papacy were the product of greed on the part of individual popes or a changed economic climate as a result of increased secular power is a topic of debate that this paper does not seek to answer. This paper refers to such fiscal policies as "abuses" because they were viewed as such by the reformers at the Council of Constance, who sought to legislate against them and ultimately led to the changes in policy which this paper addresses. For more in-depth discussions of the fiscal policies of the Avignon Papacy and the debate surrounding them, see, for example, Ludwig Von Pastor, *History of the Popes Vol. 1, The Great Schism*, Norman Housley, *The Avignon Papacy and the Crusades, 1305-1378* (Oxford, Clarendon Press) 1986, and Daniel Williman, "The Right of Spoil of the Popes of Avignon 1316-1415." *Transactions of the American Philosophical Society* 78, no. 6 (1988).

^BThe method of self-assessment and establishment of a lay/clerical committee only applied to England. In France, where the tax saw significantly less success, collection was carried out by individual seigneurs upon their lay tenants. The tax was resented in France, and Philip Augustus abandoned it after a year. Royal enforcement of the tax was only present in England.

^CBishops could not absolve an individual from excommunication in certain cases. This right was reserved for the Papal Curia, and the individual was meant to make a personal appeal in Rome in order to have the excommunication lifted.

^DOn the responsive nature of Honorius' pontificate, see Thomas W. Smith. *Curia and Crusade: Pope Honorius III and the Recovery of the Holy Land, 1216-1227*, Outremer: Studies in the Crusades and the Latin East 6 (Turnhout: Brepols, 2017). On resistance to taxation during the pontificates of Innocent, Honorius, and Gregory, see Jessalyn Bird. "Finance of Crusades" in A.V. Murray. *The Crusades: An Encyclopedia, Volume II* (ABC-CLIO: Santa Barbara, 2006) 442-46.

^ESee Brenda Bolton, "Hearts Not Purses: Innocent III's Attitude to Social Welfare." In *Through the Eye of a Needle: Judaeo-Christian Roots of Social Welfare*, edited by Emily Albu Hanawalt and Carter Lindberg (Missouri: The Thomas Jefferson University Press at Northeast Missouri State University, 1994) 123-145, as well as Bird, "Peter the Chanter's Circle" and Parker, "Papa et Pecunia."

^FOn the role of commutation and lay participation, see Jessalyn Bird. "Vows" in A.V. Murray. *The Crusades: An Encyclopedia, Volume IV* (ABC-CLIO: Santa Barbara, 2006) 1233-1237.

^GSpoils refer to the moveable property of deceased or intestate clerics. Annates refer to a payment to the Curia from a recipient of an ecclesiastic benefice. Services refer to the payment to the curia in return for use of the papal chancery.

^H"Levies for various crusades also contributed to the development of centralized financial administrations and the growth of papal and royal taxation, at the same time aiding the development of representative bodies

whose consent was required for many forms of taxation." Bird, "Financing the Crusades," 446.

VIII. References

¹Parker, M. (2017) "Papa et Pecunia : Innocent III's Combination of Reform and Fiscal Policy to Finance Crusades." *Mediterranean Historical Review* 32, no. Pg. 6.

²Graves, Edgar. (1935) review of Papal Revenues in the Middle Ages, by Lunt, William E, *Speculum* 10, no. 2. Pg. 218-221.

³Lunt, W. (1934) Papal Revenues in the Middle Ages. Vol. 1. Columbia. Pg. 72.

⁴Ibid, 10.

⁵Brundage, J. (1969) Medieval Canon Law and the Crusader. UMI. Pg. 187.

⁶Cazel, F. (1955) "The Tax of 1185 in Aid of the Holy Land." *Speculum* 30, no. 3. Pg. 385.

⁷Ibid, 388.

⁸Ibid, 385-389.

⁹Ibid.

¹⁰Brundage, Medieval Canon Law and the Crusader, 195-6.

¹¹Cazel, "1185 Tax", 126.

¹²Ibid.

¹³Ibid.

¹⁴Ibid.

¹⁵Ibid, 129.

¹⁶Ibid, 135.

¹⁷Ibid, 148.

¹⁸Queller, D., Andrea, A. and Madden, T. (2000) *The Fourth Crusade: the Conquest of Constantinople*. University of Pennsylvania Press. Pg. 1-3.

¹⁹Kay, R., (1980) "The Albigensian Twentieth of 1221-3: An Early Chapter in the History of Papal Taxation." *Journal of Medieval History* 6, no. 3. Pg. 308.

²⁰Brundage, Medieval Canon Law and the Crusader, 187.

²¹Powell, J., (1994) *Anatomy of a crusade: 1213-1221*. University of Pennsylvania Press. Pg. 91-92.

²²Cazel, F. (1989) "Financing the Crusades" in *The Impact of the Crusades on Europe*. Hazard, H., Zacour, N., and Setton, K., eds. University of Wisconsin Press. Pg. 137.

²³Kay, "The Albigensian Twentieth", 313.

²⁴Ibid, 310.

²⁵Ibid, 313.

²⁶France, J., (2005) *The Crusades and the Expansion of Catholic Christendom, 1000-1714*. Routledge. Pg. 106.

²⁷Collins, R., (2010) *Keepers of the Keys of Heaven: A History of the Papacy*. Phoenix. Pg. 259.

²⁸Ibid. Pg. 263.

²⁹Asbridge, T., (2012) *The Crusades: The War for the Holy Land*. Simon & Schuster. Pg. 385-387.

³⁰Bird, J. (2006) "Financing the Crusades" in Murray, A.V. *The Crusades: An Encyclopedia*, Volume II. ABC-CLIO: Santa Barbara. Pg. 442-46.

³¹Lunt, Papal Revenues, 76-7.

³²Brundage, Medieval Canon Law, 187.

³³Ibid.

³⁴Ibid.

³⁵Housley, N., (1982) *The Italian Crusades: The Papal-Angevin Alliance and the Crusades Against Christian Lay Powers, 1254-1343*. Clarendon.

³⁶Ibid, 178.

³⁷Ibid.

³⁸Ibid.

³⁹Ibid, 200-1.

⁴⁰Ibid.

⁴¹Stump, P., (1989) "The Reform of Papal Taxation at the Council of Constance (1414-1418)." *Speculum* 64, no. 1. Pg. 76.

⁴²Ibid.

⁴³Cazel, "Financing the Crusades", 129.

⁴⁴Ibid, 149.

⁴⁵O'Callaghan, Joseph., (2013) *Reconquest and Crusade in Medieval Spain*. University of Pennsylvania

Press. Pg. 122.

⁴⁶Ibid.

⁴⁷France, *The Crusades and Christendom*, 303.

⁴⁸Brundage, *Medieval Canon Law*, 135-6; Lunt, *Papal Revenues*, 115.

⁴⁹Riley-Smith, J. (1986) *The First Crusade and the Idea of Crusading*. University of Pennsylvania Press. Pg. 43-47.

⁵⁰Ibid.

⁵¹Ibid.

⁵²Ibid.

⁵³Smith, T.W. (2019). "How to Plan a Crusade Call: Pope Innocent III and *Quia maior* (1213)." *Historical Research*, volume 92, issue 255. Pg. 7-8.

⁵⁴France, *The Crusades and Christendom*, 160-163.

⁵⁵Ibid.

⁵⁶Lower, M., (2005) *The Barons Crusade: A Call to Arms and Its Consequences*. University of Pennsylvania Press. Pg. 16.

⁵⁷Ibid.

⁵⁸Bird, J. (2003). "Innocent III, Peter the Chanter's Circle, and the Crusade Indulgence: Theory, Implementation, and Aftermath." In *Innocenzo III: Urbs et orbis, atti del Congresso Internazionale* (Roma, 9-15 settembre 1998) [*Innocent III: City and World, Acts of the International Congress (Rome 9-15 September 1998)*], edited by Andrea Sommerlechner, vol 1. Rome: Società romana di storia patria. Pg. 509.

⁵⁹Parker, "Papa et Pecunia". Pg. 7; Bird, "Peter the Chanter's Circle", Pg. 509.

⁶⁰Ibid, 10.

⁶¹Bird, "Peter the Chanter's Circle," 509. Parker, "Papa et Pecunia", 12.

⁶²Bird, "Peter the Chanter's Circle", 509.

⁶³Lunt, *Papal Revenues*, 112.

⁶⁴Ibid.

⁶⁵Brundage, *Medieval Canon Law*, 75.

⁶⁶Hübner, B., (1867) *Die Constanzer Reformation und die Concordate von 1418* B. Tauchnitz, Pg. 360-388; Valois, N., (1896) *La France et le Grand Schisme d'Occident*, Vol. 4. A. Picard. Pg. 419-426.

⁶⁷Lunt, *Papal Revenues*, 93-9, 103-7, 126.

⁶⁸Hübner, *Die Constanzer Reformation und die Concordate von 1418*, 360-88.

⁶⁹Stump, "Reform of Papal Taxation", 104.

⁷⁰Ibid, 73.

⁷¹Ibid.

⁷²Ibid, 72.

⁷³Stump, "Reform of Papal Taxation," 98.

⁷⁴Ibid, 72.

⁷⁵Ibid, 105.

⁷⁶Collins, *Keeper of the Keys of Heaven*, 350.

⁷⁷Collins, 6. See also Phillips, Jonathan. *The Second Crusade: Extending the Frontiers of Christendom* (New Haven: Yale University Press, 2010). Pg. 106.

⁷⁸Parker, "Papa et Pecunia", 6.

⁷⁹Ibid. See also Bird, "Peter the Chanter's Circle," 506-8.

⁸⁰Smith, T.W. (2017). *Curia and Crusade: Pope Honorius III and the Recovery of the Holy Land, 1216-1227*, Outremer: Studies in the Crusades and the Latin East 6. Turnhout: Brepols.

